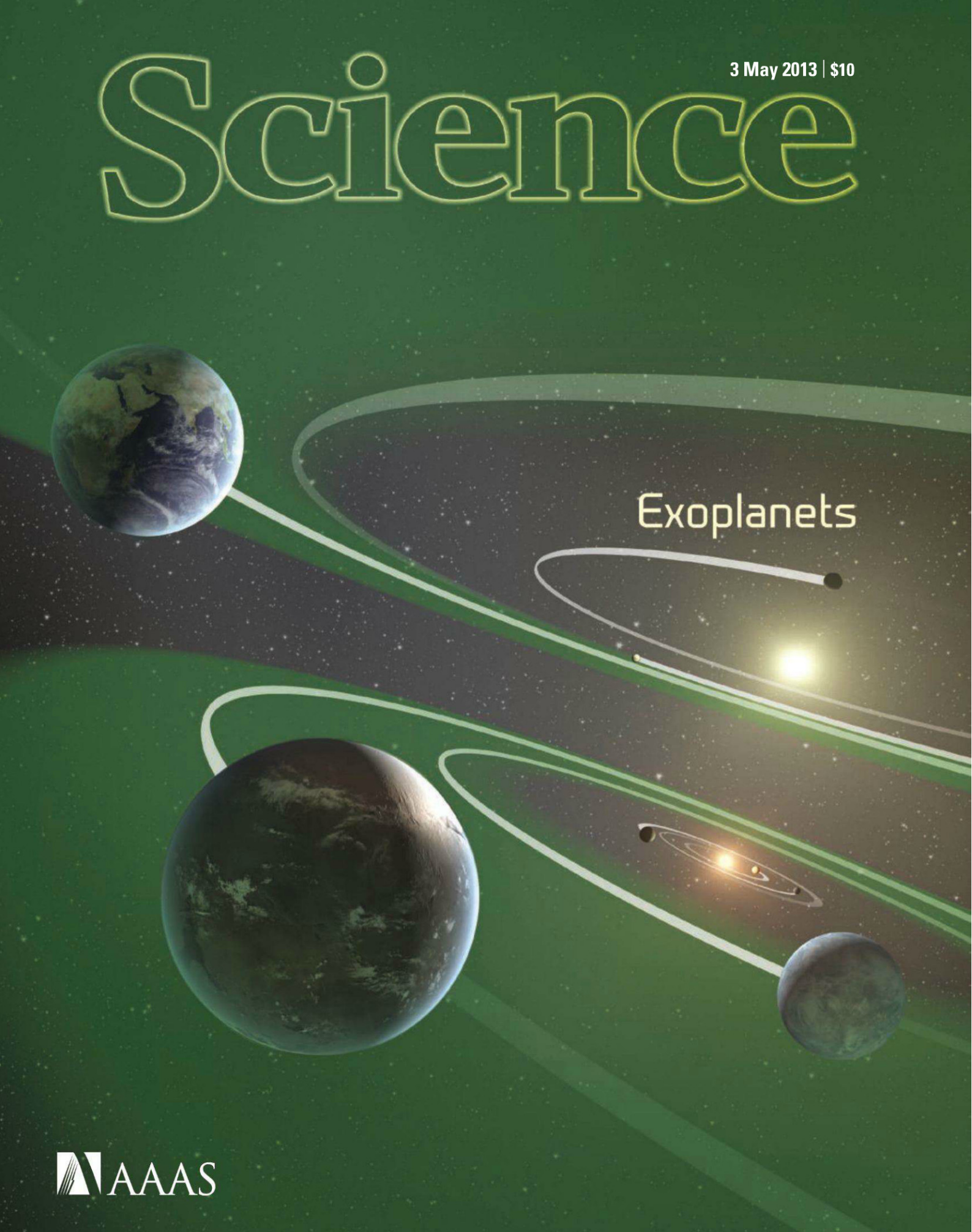
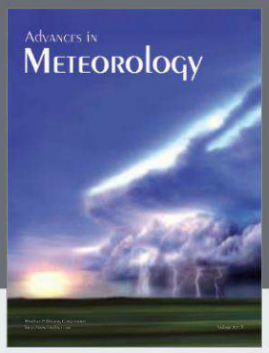
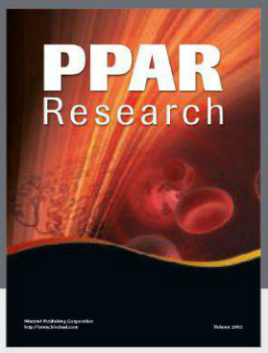
Science

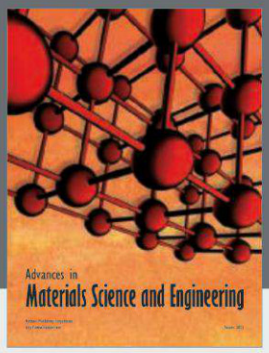
3 May 2013 | $10

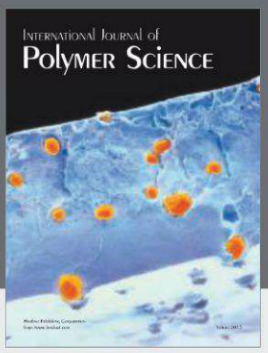


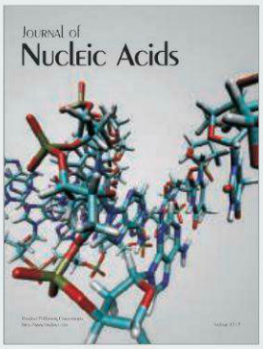
AAAS

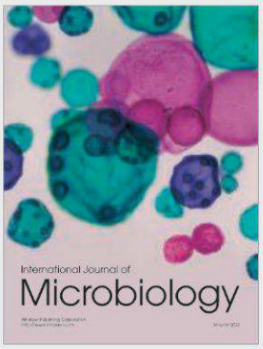






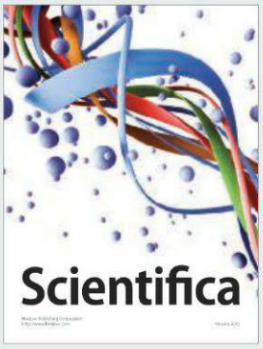


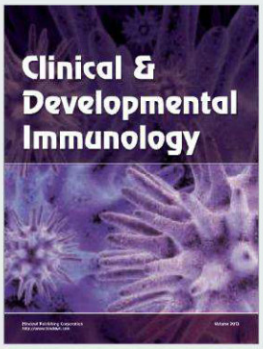


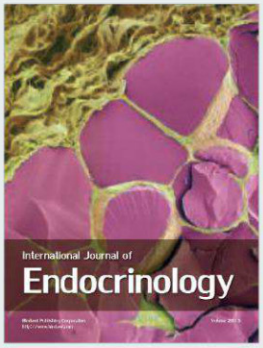


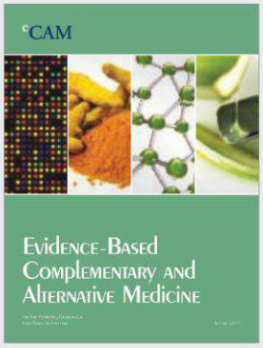
Hindawi

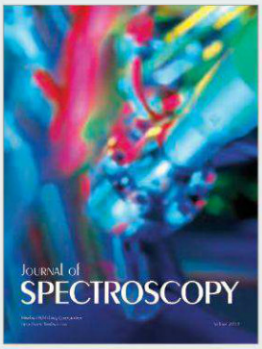
Submit your manuscripts at
http://www.hindawi.com

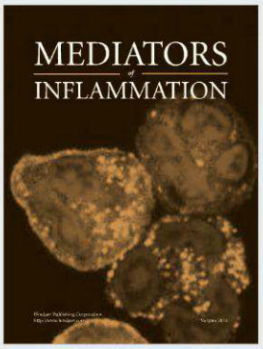


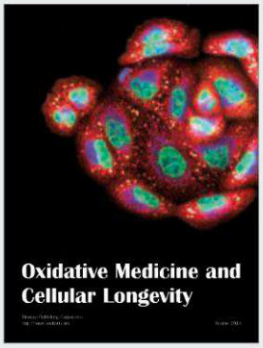


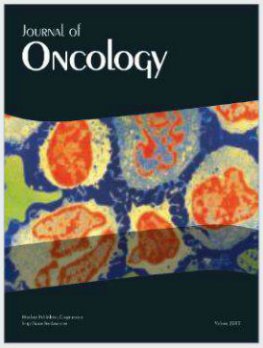


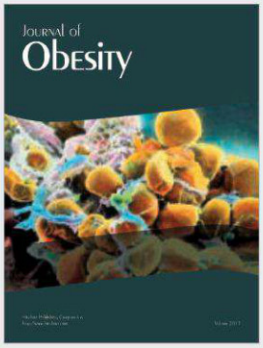


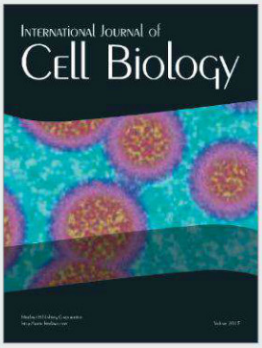


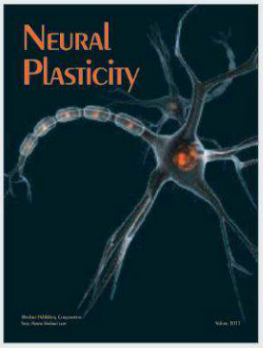




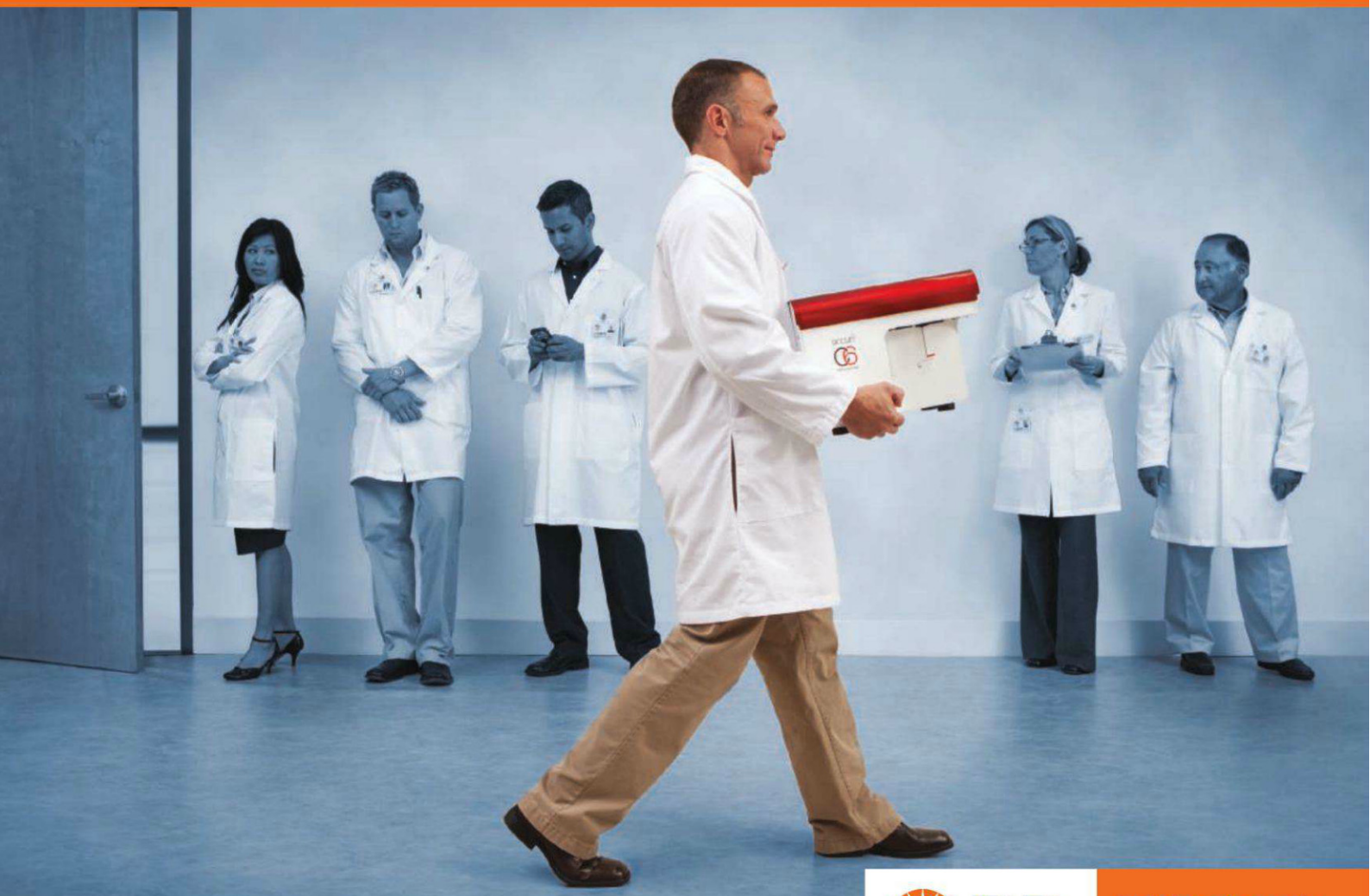








The BD Accuri™ personal flow cytometer



BD

Helping all people live healthy lives

Flow cytometry within reach.™

The BD Accuri™ Cytometer is a personal flow cytometer that puts flow cytometry within reach. It gives you 4-color cell analysis in an affordable, transportable, and easy-to-use format.

Measuring just 11 x 14.75 x 16.5 inches (27.9 x 37.5 x 41.9 centimeters) and weighing just 30 pounds (13.6 kilograms), the BD Accuri Cytometer gives both novice and experienced researchers the power of multicolor analysis, when and where you need it. The software's intuitive interface guides you through workflows, making it easy to begin collecting and analyzing data—even if you have little flow cytometry knowledge or know-how.

Setup and maintenance are also simplified to increase availability and up-time. In the event service is required, just pack up the system and return it. A loaner system will be provided while your system is serviced.

For more information about how you can more easily access the power of flow cytometry in your lab, visit **bdbiosciences.com/go/accuri**.

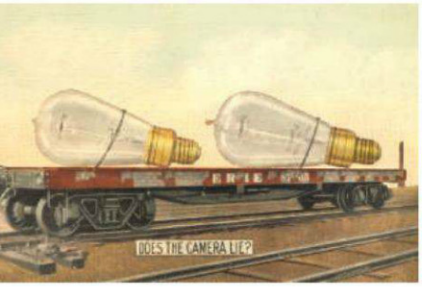
Flow cytometry within reach.™

BD flow cytometers are Class 1 Laser Products.
For Research Use Only. Not for use in diagnostic or therapeutic procedures.
BD, BD Logo and all other trademarks are property of Becton, Dickinson and Company. © 2012 BD
23-13463-01

BD Biosciences
2350 Qume Drive
San Jose, CA 95131
bdbiosciences.com

Volume 340 Issue 6132

page 546



page 551

SPECIAL SECTION

Exoplanets

INTRODUCTION

565 Alien Worlds Galore

NEWS

566 Gallery of Planet Hunters

570 … And a Glossary of Their Quarry
>> Science *Podcast*

REVIEWS

572 Observed Properties of Extrasolar Planets
A. W. Howard

577 Exoplanet Habitability
S. Seager
>> Science *Podcast*

>> News story p. 542; Report p. 587; Planet Hunters slideshow; and more at www.sciencemag.org/special/exoplanets

EDITORIAL

525 Is Any Science Safe?
Kenneth Prewitt

NEWS OF THE WEEK

530 A roundup of the week's top stories

NEWS & ANALYSIS

534 Bill Would Set New Rules for Choosing NSF Grants

535 Research in Limbo as Harvard Moves to Close Center

536 Bill and Melinda Gates Talk Science

539 Scientists Clash Swords Over Future of GM Food Crops in India

540 Hansen's Retirement From NASA Spurs Look at His Legacy

NEWS FOCUS

542 Mr. Borucki's Lonely Road to the Light
>> Exoplanets section p. 565; Report p. 587

546 Taking the Pulse of a Ravaged Ocean
The Pacific Swallows Fukushima's Fallout

LETTERS

548 Geoengineering: Guidance Exists
P. S. Corden

Geoengineering: Perilous Particles
B. Fadeel et al.

549 Life in Science: The Legend of Owen

549 TECHNICAL COMMENT ABSTRACTS

BOOKS *ET AL.*

550 Surfaces and Essences
D. Hofstadter and E. Sander, reviewed by K. J. Holyoak

551 Faking It
M. Fineman, curator;
Faking It
M. Fineman

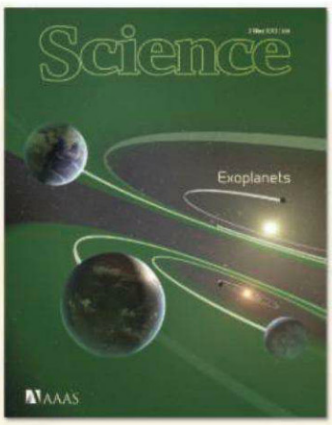
POLICY FORUM

552 Global Research Integrity Training
N. H. Steneck

ON THE WEB THIS WEEK

>> Special Issue Slideshow
Check out a gallery of planet hunters.

>> Find More Online
Check out *Science* Express, our podcast, videos, daily news, our research journals, and *Science* Careers at **www.sciencemag.org**.



COVER

Stylized comparison of the inner solar system (top) and Kepler-62 (bottom), a five-planet system about 1200 light-years away. Like our solar system, Kepler-62 hosts two planets in its habitable zone (depicted in green), the region around a star where a planet could conceivably maintain liquid water on its surface. The special section beginning on page 565 summarizes what we currently know about extrasolar planets.

Illustration: NASA/Jet Propulsion Laboratory–Caltech/T. Pyle

DEPARTMENTS

523 This Week in *Science*
526 Editors' Choice
528 *Science* Staff
643 New Products
644 *Science* Careers

PERSPECTIVES

554 Recombinatorial Logic
Y. Benenson
>> Report p. 599

555 Seeing Gravitational Waves
M. M. Kasliwal

556 Tracking Marine Pollution
J. E. Elliott and K. H. Elliott

558 Silencing a Metabolic Oncogene
J. Kim and R. J. DeBerardinis
>> Reports pp. 622 and 626

559 Unconventional Secretion, Unconventional Solutions
M. Zhang and R. Schekman

561 Obey the Peptide Assembly Rules
M. S. Ardejani and B. P. Orner
>> Report p. 595

562 As Good as Chocolate
K. Palczewski and P. D. Kiser
>> Reports pp. 610 and 615

RESEARCH ARTICLE

583 Emergence and Frustration of Magnetism with Variable-Range Interactions in a Quantum Simulator
R. Islam et al.
Coulomb interactions in a system of 16 trapped ions are used to simulate magnetism with varying degrees of frustration.

REPORTS

587 Kepler-62: A Five-Planet System with Planets of 1.4 and 1.6 Earth Radii in the Habitable Zone
W. J. Borucki et al.
The Kepler mission detected a five-planet system with two small planets in the habitable zone of a star lighter than the Sun.
>> News story p. 542; Exoplanets section p. 565

591 Complex N-Heterocycle Synthesis via Iron-Catalyzed, Direct C–H Bond Amination
E. T. Hennessy and T. A. Betley
An iron catalyst facilitates carbon–nitrogen bond formation in previously unreactive substrates.

595 Self-Assembling Cages from Coiled-Coil Peptide Modules
J. M. Fletcher et al.
Hexagonal networks form from heterodimeric and homotrimeric coiled coils and create ~100-nanometer-diameter cages.
>> Perspective p. 561

599 Amplifying Genetic Logic Gates
J. Bonnet et al.
A genetic circuit architecture resembling a transistor can be engineered into individual live cells.
>> Perspective p. 554

603 Controlled Flight of a Biologically Inspired, Insect-Scale Robot
K. Y. Ma et al.
A fly-scale, flapping-wing robot has been fabricated and can be controlled to perform unconstrained flight maneuvers.

607 3D Reconstruction of the Source and Scale of Buried Young Flood Channels on Mars
G. A. Morgan et al.
Radar data taken from orbit reveal the complex morphology of buried flood channels in a young region of Mars.

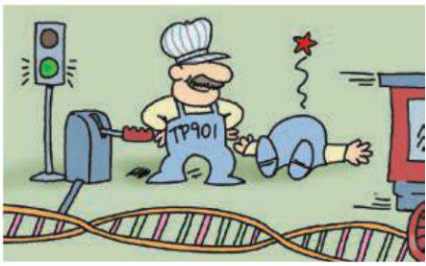
610 Structural Basis for Molecular Recognition at Serotonin Receptors
C. Wang et al.

615 Structural Features for Functional Selectivity at Serotonin Receptors
D. Wacker et al.
Structures of serotonin receptor family members in complex with the fungal alkaloid ergot offer clues for drug designers.
>> Perspective p. 562

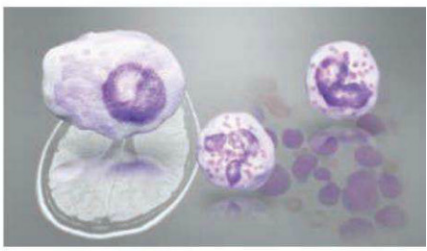
619 R-Loop Stabilization Represses Antisense Transcription at the *Arabidopsis FLC* Locus
Q. Sun et al.
A homeodomain protein limits expression of antisense transcripts involved in flowering time control.

622 Targeted Inhibition of Mutant IDH2 in Leukemia Cells Induces Cellular Differentiation
F. Wang et al.

626 An Inhibitor of Mutant IDH1 Delays Growth and Promotes Differentiation of Glioma Cells
D. Rohle et al.
A small molecule that inhibits a mutant enzyme in tumors slows malignant growth by inducing cancer cell differentiation.
>> Perspective p. 558



pages 554 & 599



pages 558, 622, & 626

630 Disparate Individual Fates Compose Robust CD8+ T Cell Immunity
V. R. Buchholz et al.

635 Heterogeneous Differentiation Patterns of Individual CD8+ T Cells
C. Gerlach et al.
The single-cell dynamics as cytotoxic T cells respond to a bacterial infection are analyzed in mice.

639 Neural Decoding of Visual Imagery During Sleep
T. Horikawa et al.
Machine-learning models can predict specific visual dream contents from brain activity measurement alone.

CREDIT: (BOTTOM) ERIC D. SMITH

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals Mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Copyright © 2013 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership and subscription (51 issues): $149 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $990; Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (air assist delivery) $85. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #1254 88122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**

Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $10.00 current issue, $15.00 back issue prepaid includes surface postage; bulk rates on request. **Authorization to photocopy** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by AAAS to libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that $30.00 per article is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

BD Horizon™ Brilliant Violet™ Polymer Conjugates

Bright dyes help resolve rare and dim cell populations

BD

Helping all people live healthy lives

Helping nature to reveal itself.

Now you can resolve rare or dim cell populations with BD Horizon™ Brilliant Violet™ polymer conjugates from BD Biosciences.

Developed from pioneering polymer dye technology acquired from Sirigen Ltd., BD Horizon Brilliant Violet dyes are brighter than conventional dyes. Improved brightness enables you to identify cell populations with lower receptor density than was previously possible, resolving cell populations previously obscured.

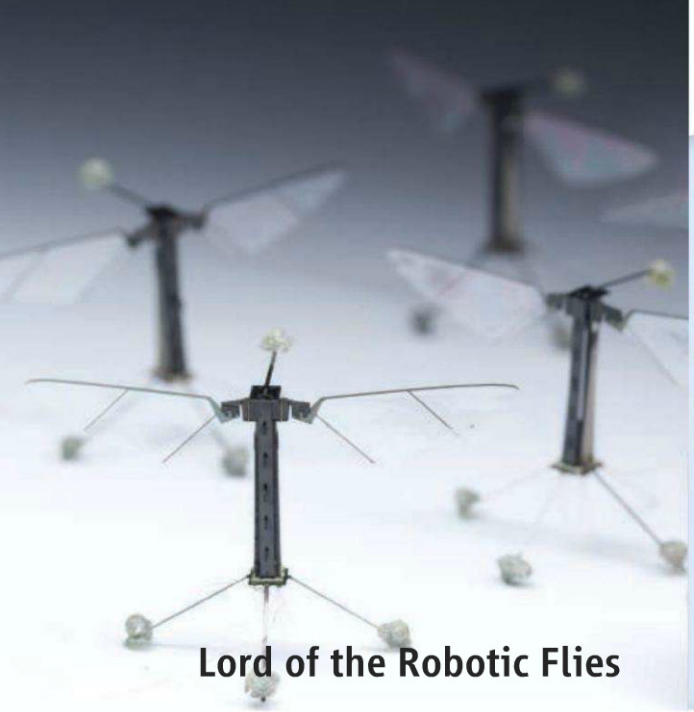
The complete portfolio of BD conjugated antibodies can be used to explore cellular features and characterize cells through surface, intracellular, or secreted markers. To ensure you can use BD reagents across your entire multicolor panel, our portfolio contains a broad selection of fluorochrome-conjugated antibodies.

Request a free sample, or find out how you can use BD Horizon Brilliant Violet polymer conjugates with our expanded set of tools and information at **bdbiosciences.com/go/brilliant.**

For Research Use Only. Not for use in diagnostic or therapeutic procedures.
BD, BD Logo and all other trademarks are property of Becton, Dickinson and Company. © 2013 BD
23-14924-00

BD Biosciences
2350 Qume Drive
San Jose, CA 95131
bdbiosciences.com

EDITED BY STELLA HURTLEY



Lord of the Robotic Flies

While small-scale flying objects are ubiquitous in nature, they are quite hard to engineer. As sizes get smaller, fixed-winged flight becomes less efficient because of increased drag resistance. **Ma *et al.*** (p. 603) developed a tethered robotic fly with wings that flap through the use of piezoelectric materials. Control of the flight motion involved a feedback process, which allowed the tethered robotic fly to hover and make controlled flight maneuvers.

Magnetic Frustration

The study of magnetic frustration has a long history in solid-state physics, but cold-atom systems now offer the possibility of simulating the problem with exquisite control. **Islam *et al.*** (p. 583) study a system of 16 trapped ions, using the Coulomb interactions between the ions to simulate exchange interactions present in magnetic systems. The measured spin correlations provide a window into the behavior of the system, as the effective magnetic field and the range of the interactions are tuned.

Closing the Cycle

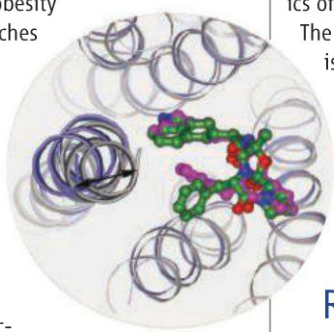
Cyclic hydrocarbons that incorporate nitrogen in the ring are among the most heavily investigated compounds in medicinal chemistry. **Hennessy and Betley** (p. 591) demonstrate an iron catalyst that forms a range of such cyclic compounds by inducing linear alkyl azides to curl back on themselves, inserting the nitrogen at one end into a carbon-hydrogen bond further down the chain. The reaction furthers a trend of C-H bond activation chemistry that forms elaborate products from relatively simple precursors, without the need to install activating groups at unreactive sites.

Seeing Below the Surface

Marte Vallis represents the largest of the young outflow channels on Mars. Thought to be the result of an ancient megaflood that occurred during a period that is otherwise known to be cold and dry, this channel system extends more than approximately 1000 kilometers in length and 100 kilometers in width. The channels' depths are unknown, however, because they have been covered by lava flows. By using the SHARAD radar sounder on the Mars Reconnaissance Orbiter, **Morgan *et al.*** (p. 607, published online 7 March) produced a three-dimensional reconstruction of the subsurface channels. The channels are twice as deep as previously thought and the floodwaters came from the Cerberus Fossae extensional fracture system.

Dissecting Serotonin Receptors

Serotonin receptors are the targets for many widely used drugs prescribed to treat ailments from depression to obesity and migraine headaches (see the Perspective by **Palczewski and Kiser**). **C. Wang *et al.*** (p. 610, published online 21 March) and **Wacker *et al.*** (p. 615, published online 21 March) describe crystal structures of two members of the serotonin family of receptors bound to antimigraine medications or to a precursor of the hallucinogenic drug LSD. Subtle differences in the way particular ligands bind to the receptors cause substantial differences in the signals generated by the receptor and the consequent biological responses. The structures reveal how the same ligand can activate one or both of the two main serotonin receptor signaling mechanisms, depending on which particular receptor it binds.



IDHology

Among the most exciting drug targets to emerge from cancer genome sequencing projects are two related metabolic enzymes, isocitrate dehydrogenases 1 and 2 (IDH1, IDH2). Mutations in the *IDH1* and *IDH2* genes are common in certain types of human cancer. Whether inhibition of mutant IDH activity might offer therapeutic benefits is unclear (see the Perspective by **Kim and DeBerardinis**). **F. Wang *et al.*** (p. 622, published online 4 April) isolated a small molecule that selectively inhibits mutant IDH2, describe the structural details of its binding to the mutant enzyme, and show that this compound suppresses the growth of patient-derived leukemia cells harboring the *IDH2* mutation. **Rohle *et al.*** (p. 626, published online 4 April) show that a small molecule inhibitor of IDH1 selectively slows the growth of patient-derived brain tumor cells with the *IDH1* mutation.

Dynamic Protection

During an immune response, $CD8^+$ T cells are recruited to provide protection. Most cells differentiate into short-lived effectors that help to clear the pathogen, whereas others form long-lived memory cells to protect against reinfection. **Gerlach *et al.*** (p. 635, published online 14 March) and **Buchholz *et al.*** (p. 630, published online 14 March) used in vivo fate mapping of mouse T cells with a defined specificity during a bacterial infection to show that the dynamics of the single-cell response are not uniform. The response of a particular T cell population is the average of a small number of clones that expand greatly ("large clones") and many clones that only proliferate at low amounts ("small clones"). The memory pool arises largely from small clones whereas effectors are primarily made up of large clones.

Reading Dreams

How specific visual dream contents are represented by brain activity is unclear. Machine-learning–based analyses can decode the stimulus- and task-induced brain activity patterns that represent specific visual contents. **Horikawa *et al.*** (p. 639, published online 4 April) examined patterns of brain activity during dreaming and compared these to waking responses to visual stimuli. The findings suggest that the visual content of dreams is represented by the same neural substrate as observed during awake perception.

CREDITS (TOP TO BOTTOM): KEVIN Y. MA, PAKPONG CHIRARATTANANON; C. WANG *ET AL.*

Do you have a *question* with the potential to change the world?

Do you have an idea for a research question that requires the attention of some of the world's best minds – a question whose ambitious scope and depth will spark transformative inquiry? This is your opportunity to create an interdisciplinary global research network.

The Canadian Institute for Advanced Research invites proposals from individuals or small teams to establish a research program that addresses a complex, fundamental question of importance to humanity. Areas of inquiry can draw on disciplinary expertise anywhere across the natural sciences, health and biological sciences, social sciences and humanities. CIFAR will support successful candidates in sustaining a collaborative network for a five-year renewable term.

Established in 1982, CIFAR brings together some of the brightest minds in Canada and around the world to create knowledge that enriches human life and leads to a better future.

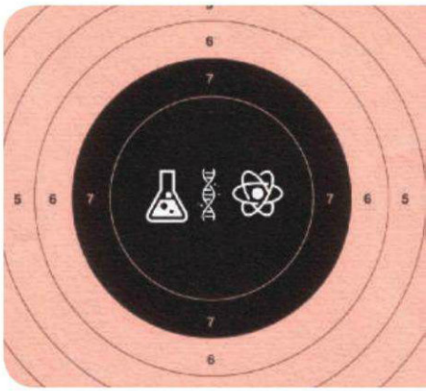
CIFAR GLOBAL CALL FOR IDEAS
CLOSES JUNE 7, 2013
cifar.ca/globalcall

CIFAR
CANADIAN INSTITUTE FOR ADVANCED RESEARCH

Kenneth Prewitt is the Carnegie Professor at the School of International and Public Affairs and director of the Scholarly Knowledge Project at Columbia University, New York. E-mail: kp2058@columbia.edu.

Is Any Science Safe?

THIS MONTH, THE U.S. NATIONAL SCIENCE FOUNDATION (NSF) WILL APPLY TWO CRITERIA IN ITS review of research proposals: intellectual merit and impact. One discipline, however, will have to meet a further test. In March, at the urging of U.S. Senator Tom Coburn (R-OK), Congress halted funding for political science—"except for" research that the agency's director certifies as "promoting national security or the economic interests of the United States." This extra test might not stop with political science. Representative Bill Posey (R-FL), in an NSF oversight hearing on 17 April, asked John Holdren, President Obama's science adviser, why Coburn's two criteria were not "a good and proper filter" to apply to all NSF grants, eliciting this response: ". . . it's a dangerous thing for Congress, or anybody else, to be trying to specify in detail what types of fundamental research NSF should be funding."* The United States has benefitted enormously from a government/science partnership carefully designed by Congress over many decades. There have been bumps along the way; it is challenging to reconcile scientific autonomy with Congress's responsibility for public funds. Coburn's "except for" clause is potentially a very large bump.



This congressional attempt to micromanage NSF grants carries three risks. First, it favors research that promises near-term benefits, overlooking the fact that there is knowledge useful under today's conditions and knowledge that becomes useful when conditions change. In the 1930s, political scientists, historians, and economists working in China and Japan were of little use to the U.S. government. But early in World War II, their expertise on the Far East was suddenly in great demand at the Office of Strategic Services, America's first intelligence agency. With few exceptions, Congress has supported both present- and future-oriented research. One well-known exception was the 1970 Mansfield amendment, which restricted the Department of Defense (DOD) to research narrowly targeted to specific military missions. Had this restriction not been lifted, DOD research in the 1980s that led to the Internet would have gone unfunded. Today, we cannot know how and when the science of the Higgs boson subatomic particle will prove useful. But conditions will change; the knowledge will be used.

The second risk is that Coburn's criteria weaken the way science builds theories, without which there is no scientific explanation of anything. Coburn recognizes that political science contributes to an understanding of national security and the economy. He misses that this understanding is embedded in broad theories about how governments work, which in turn involves studying topics seemingly unrelated to security or the economy: bureaucratic inefficiencies, moral hazards, unintended consequences, organizational decision-making, coalition-building, and much more. Science is an interconnected enterprise. Research on schoolyard bullies can unexpectedly lead to theory that explains suicide bombers. Two U.S. political scientists, Herbert Simon and Elinor Ostrom, received Nobel Prizes for theoretical work on government decision-making under uncertainty. Their theories are broadly applicable, including in explanations of failed states—often home to terrorist cells.

The third risk is to peer review. Congressional intimidation lurks in legislation that instructs the NSF director to certify individual grants. This invites responsiveness to perceived congressional priorities rather than reliance on the search for excellence through peer review. The risk of marginalizing peer review is especially worrisome given the already insecure status of politically contested science, such as evolution, stem cells, climate change, and alternative energies. Members of Congress who believe that the executive branch should not try to pick winners and losers in the market economy should certainly realize that the legislative branch should not try to pick winners and losers in science. Every scientific discipline has a stake in undoing the damage inflicted on political science, and, in fact, to the national interest, by the Coburn criteria. Every scientist should vigorously contest any effort to apply those criteria more broadly.

– Kenneth Prewitt

10.1126/science.1239180

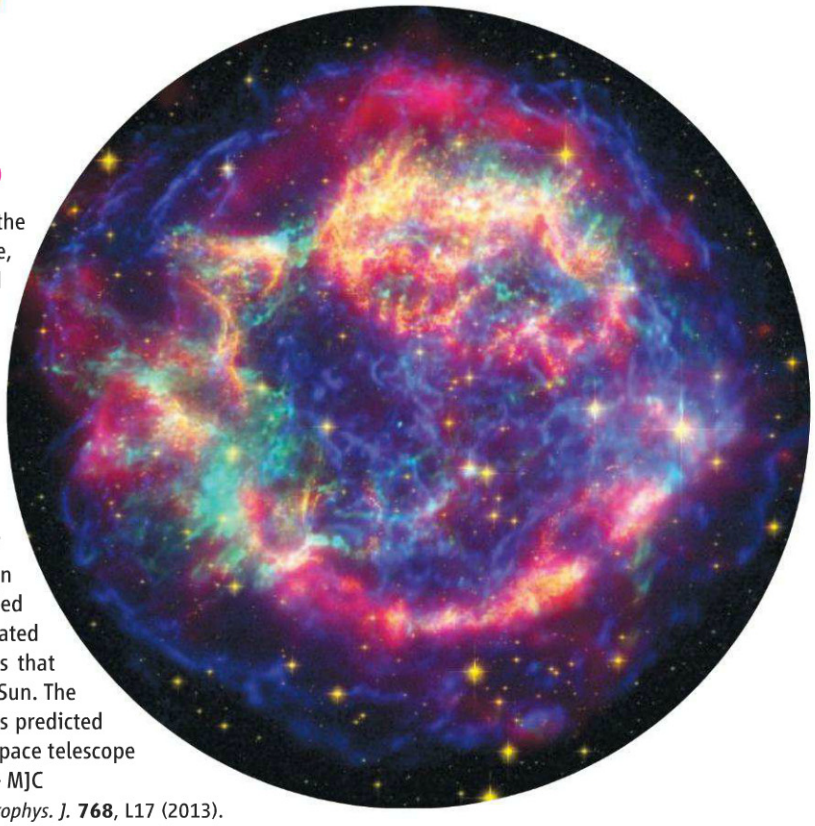
*J. Mervis, http://news.sciencemag.org/scienceinsider/2013/04/nsf-peer-review-under-scrutiny-b.html (2013).

CREDITS: (TOP) SUSAN VOGEL/PRINCE STREET PICTURES; (RIGHT) ISTOCKPHOTO

EDITED BY MARIA CRUZ AND SACHA VIGNIERI

GEOCHEMISTRY

Supernova Grains Identified in the Lab



Primitive meteorites contain grains of materials that predate the formation of the solar system. These materials—which include, among others, diamond, graphite, and various oxides and silicates—were present in the molecular cloud from which the Sun and the solar system formed and were later incorporated into primitive solar system solids. Presolar grains are identified by their unusual isotopic compositions, which can only be explained if they formed outside the solar system, in the outflows or explosions of other stars. Using nano secondary ion mass spectrometry (NanoSIMS) and Auger electron spectroscopy, Haenecour *et al.* report the identification of two silica (SiO_2) grains in the primitive meteorites LaPaz 031117 and Grove Mountains 021710. The oxygen isotopic composition of these two grains is different from those of previously identified presolar SiO_2 grains, and it indicates that the two grains originated from core collapse supernovae: the thermonuclear explosions that end the lives of stars heavier than eight times the mass of the Sun. The presence of SiO_2 dust in matter ejected by such explosions was predicted by theoretical models; recent observations by NASA's Spitzer Space telescope have also suggested their presence in supernova remnants. — MJC

Astrophys. J. **768**, L17 (2013).

ECOLOGY

Palm Monoculture Bad for Birds

The conversion of tropical forest to oil palm plantations has rapidly increased over the past decade, predominantly in Southeast Asia, where such cultivation now dominates over 2 million hectares. Substantial biodiversity loss accompanies such conversion, but little is known of the ecology of the resulting landscape. Azhar *et al.*'s survey of bird faunas in plantations and logged swamp forest in Malaysia shows that guilds were affected in different ways. Notably, raptors were more abundant in plantations than in logged forest, whereas the reverse was true for insectivores and granivores. Patterns within plantations were also influenced by the management regime (e.g., smallholding versus estate) and proximity to forest. Edwards *et al.* surveyed the functional diversity—a measure incorporating foraging, morphology, and behavior—of bird faunas across habitat gradients (from plantation to logged and primary forests) in Borneo. Functional diversity was similar between logged and primary forest but greatly reduced in plantations, with just a few generalist species filling a wide range of functional roles. These studies demonstrate that continued conversion from logged forest to oil palm plantation will lead to further losses of species and functional diversity. — AMS

Ibis **155**, 297; 313 (2013).

EDUCATION

Gaming Knowledge

Video games have great potential to support educational objectives. However, problems with video games in an educational setting, including use of classroom time and limited connections to standardized tests, persist. Sadler *et al.* examined the extent to which high-school students, at general, honors, and advanced levels, learned biological concepts through the use of Mission Biotech (MB), a game designed to support education in genetics and molecular biology through classroom integration with a supporting curriculum aligned to state science standards. Ten teachers were trained to use MB in a classroom setting and were asked to complete a daily activity log and postimplementation survey. Supporting instructional activities, also aligned to state standards yet independent from the MB interface, were provided. Assessments were designed in coordination with each of MB's four levels and were embedded within the game environment. Results from pre- and post-tests show statistically significant gains in student performance, on both curriculum-aligned and standards-aligned exams. Students at all levels showed gains in knowledge, but the increase was highest among the general-level students. As the United States moves toward implementation of the inquiry-based Next Generation Science Standards, evidence of game-based curriculum supporting student learning increases the potential that video games may have in enhancing science education. — MM

J. Res. Sci. Teach. **50**, 479 (2013).

PHYSICS

Entangling Spin and Light

There are several different approaches to quantum information processing currently under development, including (but not exhaustively) superconducting circuits, nitrogen vacancy defects in diamond, and semiconductor quantum dots. Each one relies on the basic building

CREDITS (TOP TO BOTTOM): NASA/JPL-CALTECH/O. KRAUSE/STEWARD OBSERVATORY; A. RABIN/FLICKR

block of a qubit, a two-level quantum system, on which the information can be stored and manipulated. In the semiconductor quantum dot approach, manipulating the spin of a single electron is promising because it has all the requirements: being fast, controllable and stable. A downside, however, has been the issue of scalability. Any useful quantum processor will require a system of many qubits, but getting quantum dots to talk to each other has presented a challenge. Addressing a preliminary point of this issue, Schaibley *et al*. show that the light (a single photon) emitted from the excited state of a charged quantum dot is entangled with the spin of the electron on the dot. The authors argue that by making the single photons emitted from other distinct quantum dots interact, it may be possible to then entangle the spins on these quantum dots and thus realize a scalable architecture. — ISO

Phys. Rev. Lett. **110**, 167401 (2013).

IMMUNOLOGY

Systems Vaccinology

Vaccination is a powerful approach for disease prevention, but why are some vaccines so effective whereas others fail? Obermoser *et al*. begin to answer this question by characterizing the immune response induced in people given either the trivalent seasonal influenza vaccine or the pneumococcal vaccine, both of which induce a protective antibody response. Gene expression profiling combined with functional assays at multiple time points after vaccination revealed that the two vaccines take different routes early, despite a similar ability to induce protective antibodies. The early response to the influenza vaccine was characterized by the induction of an antiviral and, to a lesser extent, inflammatory gene response. In contrast, the pneumococcal vaccine elicited a predominately inflammatory gene signature. By day 7 after vaccination, a developing antibody response was apparent in all vaccines. The figures are available in an interactive format, where readers can access the vast amounts of underlying data, providing an important resource for the vaccine community. — KLM

Immunity **38**, 831 (2013).

MOLECULAR BIOLOGY

Fried to a CRISPR

The CRISPR-Cas system, found in many bacteria and archea, is known as an adaptive defense against the invasion of foreign DNA. The bacteria/archea capture fragments of plasmids and phages and incorporate them into one or more CRIPSR loci. The captured sequences are then transcribed and cleaved into short CRISPR (cr)RNAs, which, when bound by a Cas nuclease complex, are used to target the invading nucleic acid for destruction.

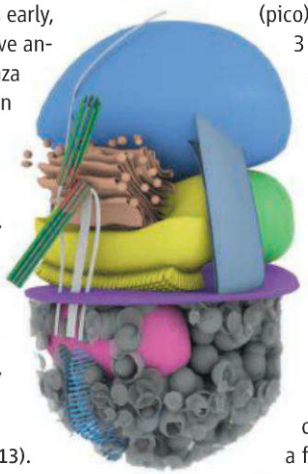
Previous work suggested a link between bacterial virulence and the CRISPR system. Sampson *et al*. now show that in the intracellular parasite *Francisella novicida,* CRISPR-Cas is required for evasion of the host's innate immune system. The *F. novicida* CRISPR locus encodes a small CRISPR/Cas-associated RNA (scaRNA) that has complementarity to both a trans-acting crRNA (tracrRNA) and its own lipoprotein gene. The scaRNA, together with the tracrRNA and Cas9 nuclease, acts to repress expression of the *F. novicida* lipoprotein during host infection. This allows the infecting bacteria to avoid detection by one of the host's pathogen pattern recognition receptors, which binds to bacterial lipoproteins, and activates the host's innate immune response. CRISPR-Cas regulation of endogenous genes may play a broad role in bacterial virulence. — GR

Nature 10.1038/nature12048 (2013).

EVOLUTION

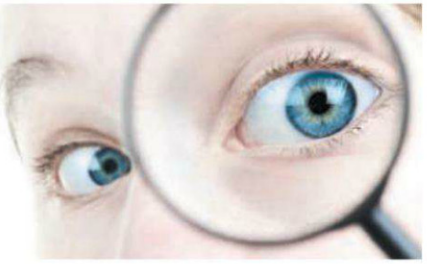
Pico- Pico- Pico-zoa Picozoa

It's not often that a new phylum is described, but the vanishingly small eukaryotes that constitute the equivalent of dark matter of the plankton have now been named the Picozoa. These highly diverse organisms, previously known as (pico)biliphytes, are about 2 to 3 μm in size and dominate aquatic ecosystems. Characterized by the pigment phycobilin and once thought to be photosynthetic, they appear instead to lack plastids and seem to be heterotrophic. Seenivasan *et al*. used a mitochondrial marker to isolate single (pico)billiphyte cells by fluorescence-activated cell sorting and established a first culture. The result was *Picomonas judraskeda,* a colloid feeder that moves through the North Sea by means of flagellae. This organism is a bipartite cell, joined at the mitochondrion, packaging all the feeding apparatus into one sac and the major organelles with the motility apparatus in the other. — CA



CREDIT: SEENIVASAN *ET AL.*, *PLoS ONE* **8**, 3 (26 MARCH 2013)

PLoS One **8**, 10.1371/journal.pone.0059565 (2013).



Still not looking at microRNA?

Try **miRSearch**

The fast track to microRNA discovery

1. Type in your favorite gene, disease or tissue
2. Find relevant microRNAs with references
3. Order qPCR primers, ISH probes or inhibitors for your microRNA

Start your **miRSearch**

exiqon.com/mirsearch

EXIQON
Seek Find Verify

www.sciencemag.org

Science

1200 New York Avenue, NW
Washington, DC 20005
Editorial: 202-326-6550, FAX 202-289-7562
News: 202-326-6591, FAX 202-371-9227
Bateman House, 82-88 Hills Road
Cambridge, UK CB2 1LQ
+44 (0) 1223 326500, FAX +44 (0) 1223 326501

Subscription Services For change of address, missing issues, new orders and renewals, and payment questions: 866-434-AAAS (2227) or 202-326-6417, FAX 202-842-1065. Mailing addresses: AAAS, P.O. Box 96178, Washington, DC 20090-6178 or AAAS Member Services, 1200 New York Avenue, NW, Washington, DC 20005

Institutional Site Licenses please call 202-326-6755 for any questions or information

Reprints: Author Inquiries 800-635-7181
Commercial Inquiries 803-359-4578

Permissions 202-326-7074, FAX 202-682-0816

Member Benefits AAAS Travels: Betchart Expeditions 800-252-4910; Apple Store www.store.apple.com/us/go/eppstore/aaas; NASA Federal, 1-888-NASA-FCU (1-888-627-2328) or www.nasafcu.com; Cold Spring Harbor Laboratory Press Publications www.cshlpress.com/affiliates/aaas.htm; GEICO Auto Insurance www.geico.com/landingpage/go51.htm?logo=17624; Hertz 800-654-2200 CDP#343457; Office Depot https://bsd.officedepot.com/portalLogin.do; Seabury & Smith Life Insurance 800-424-9883; Subaru VIP Program 202-326-6417; VIP Moving Services www.vipmayflower.com/domestic/index.html; Other Benefits: AAAS Member Services 202-326-6417 or www.aaasmember.org.

science_editors@aaas.org (for general editorial queries)
science_letters@aaas.org (for queries about letters)
science_reviews@aaas.org (for returning manuscript reviews)
science_bookrevs@aaas.org (for book review queries)

Published by the American Association for the Advancement of Science (AAAS), *Science* serves its readers as a forum for the presentation and discussion of important issues related to the advancement of science, including the presentation of minority or conflicting points of view, rather than by publishing only material on which a consensus has been reached. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.

AAAS was founded in 1848 and incorporated in 1874. Its mission is to advance science, engineering, and innovation throughout the world for the benefit of all people. The goals of the association are to: enhance communication among scientists, engineers, and the public; promote and defend the integrity of science and its use; strengthen support for the science and technology enterprise; provide a voice for science on societal issues; promote the responsible use of science in public policy; strengthen and diversify the science and technology workforce; foster education in science and technology for everyone; increase public engagement with science and technology; and advance international cooperation in science.

Information for Authors
See pages 716 and 717 of the 8 February 2013 issue or access www.sciencemag.org/about/authors

EDITOR-IN-CHIEF **Bruce Alberts**
EXECUTIVE EDITOR **Monica M. Bradford**
NEWS EDITOR **Colin Norman**
MANAGING EDITOR, RESEARCH JOURNALS **Katrina L. Kelner**
DEPUTY EDITORS **Barbara R. Jasny, Andrew M. Sugden, Valda J. Vinson**

Editorial senior editors/commentary Lisa D. Chong, Brad Wible; **senior editors** Gilbert J. Chin, Pamela J. Hines, Paula A. Kiberstis (Boston), Marc S. Lavine (Toronto), Beverly A. Purnell, L. Bryan Ray, Guy Riddihough, H. Jesse Smith, Phillip D. Szuromi (Tennessee), Jake S. Yeston, Laura M. Zahn (San Diego); **associate editors** Melissa R. McCartney (Education Projects), Kristen L. Mueller, Jelena Stajic, Sacha Vignieri (Oregon), Nicholas S. Wigginton (Michigan); **book review editor** Sherman J. Suter; **associate letters editor** Jennifer Sills; **editorial manager** Cara Tate; **senior copy editors** Jeffrey E. Cook, Cynthia Howe, Harry Jach, Lauren Kmec, Barbara P. Ordway, Trista Wagoner; **copy editor** Chris Filiatreau; **senior editorial coordinators** Carolyn Kyle, Beverly Shields; **editorial coordinators** Joi S. Granger, Anita Wynn; **publications assistants** Ramatoulaye Diop, Aneera Dobbins, Jeffrey Hearn, Lisa Johnson, Dona Mathieu, Le-Toya Mayne Flood, Shannon McMahon, Scott Miller, Jerry Richardson, Teresa R. Sakon, Brian White; **editorial assistant** Patricia M. Moore; **executive editorial assistant** Yolanda O'Bannon (San Francisco); **executive assistant** Alison Crawford; **administrative support** Maryrose Madrid

Editorial director, web and new media Stewart Wills; **senior web editor** Sarah Crespi; **web editor** Kerry Klein; **web development manager** Martyn Green; **web developer** Corinna Cohn

News deputy news editors Robert Coontz, Elizabeth Culotta, David Grimm (Online), Eliot Marshall, Jeffrey Mervis, Leslie Roberts, Richard Stone, John Travis; **contributing editor** Polly Shulman; **news writers** Yudhijit Bhattacharjee, Adrian Cho, Jennifer Couzin-Frankel, Carolyn Gramling, Jocelyn Kaiser, Richard A. Kerr, David Malakoff, Elizabeth Pennisi, Robert F. Service (Pacific NW), Erik Stokstad, Emily Underwood; **web developer** Daniel Berger; **social media strategist** Meghna Sachdev; **intern** Lizzie Wade; **contributing correspondents** John Bohannon, Jon Cohen (San Diego, CA), Ann Gibbons, Sam Kean, Eli Kintisch, Andrew Lawler, Mitch Leslie, Charles C. Mann, Virginia Morell, Gary Taubes; **copy editors** Melissa Raimondi, Kara Estelle; **administrative support** Scherraine Mack; **bureaus** San Diego, CA: 760-942-3252, FAX 760-942-4979; Pacific Northwest: 503-963-1940

Production director Wendy K. Shank; **assistant manager** Rebecca Doshi; **senior specialists** Steve Forrester, Christopher Redwood, Anthony Rosen; **Preflight director** David M. Tompkins; **manager** Marcus Spiegler; **specialists** Jason Hillman, Tara Kelly

Art director Yael Fitzpatrick; **associate art director** Laura Creveling; **senior illustrators** Chris Bickel, Katharine Sutliff; **illustrator** Yana Hammond; **senior art associates** Holly Bishop, Preston Huey; **art associates** Kay Engman, Garvin Grullón, Chrystal Smith; **photo editor** Leslie Blizard

***Science* International**

Europe (science@science-int.co.uk) **editorial: international managing editor** Andrew M. Sugden; **senior editor/commentary** Julia Fahrenkamp-Uppenbrink; **senior editors** Caroline Ash, Stella M. Hurtley, Ian S. Osborne, Peter Stern; **associate editor** Maria Cruz; **contributing editor** Helen Pickersgill; **editorial support** Rachel Roberts, Alice Whaley; **administrative support** Janet Clements, Michelle Pegrume, John Wood; **news: deputy news editor, u.k.** Daniel Clery; **contributing editor, europe** Martin Enserink; **contributing correspondents** Michael Balter (Paris), Kai Kupferschmidt (Berlin), Gretchen Vogel (Berlin)

Asia Japan Office: Asca Corporation, Tomoko Furusawa, Rustic Bldg. 7F, 77 Tenjin-cho, Shinjuku-ku, Tokyo 162-0808, Japan; +81 3 6802 4616, FAX +81 3 6802 4615, inquiry@sciencemag.jp; **contributing editor, asia** Mara Hvistendahl [China: mhvisten@aaas.org]; **contributing correspondents** Dennis Normile [Japan: +81 (0) 3 3391 0630, FAX +81 (0) 3 5936 3531; dnormile@gol.com]; Hao Xin [China: cindyhao@gmail.com]; Pallava Bagla [South Asia: +91 (0) 11 2271 2896; pbagla@vsnl.com]

EXECUTIVE PUBLISHER **Alan I. Leshner**
PUBLISHER **Beth Rosner**

Fulfillment Systems and Operations (membership@aaas.org); **customer service supervisor** Pat Butler; **specialists** LaToya Casteel, Michelle Ofordire, April Marshall; **manager, data entry** Mickie Napoleoni; **data entry specialists** JJ Regan, Jaimee Wise, Fiona Giblin

Business Operations and Administration director Deborah Rivera-Wienhold; **business systems and financial analysis director** Randy Yi; **systems analyst** Nicole Mehmedovich; **manager, business analysis** Eric Knott; **manager, business operations** Jessica Tierney; **business analysts** Cory Lipman, Celeste Troxler; **financial analyst** Jeremy Clay; **rights and permissions: administrator** Emilie David; **associate** Elizabeth Sandler; **marketing director** Ian King; **marketing managers** Alison Chandler, Julianne Wielga, Justin Sawyers; **marketing associates** Mary Ellen Crowley, Elizabeth Sattler, Rebecca Riffkin; **senior marketing executive** Jennifer Reeves; **director, site licensing** Tom Ryan; **director, corporate relations** Eileen Bernadette Moran; **senior publisher relations specialist** Kiki Forsythe; **publisher relations manager** Catherine Holland; **publisher relations, eastern region** Keith Layson; **publisher relations, western region** Ryan Rexroth; **customer relations manager** Iquo Edim; **customer relations analysts** Simon Chong, Lana Guz; **marketing manager** Christina Schlecht; **marketing associates** Paulina Curto, Mitchell Edmund; **electronic media: director** Lizabeth Harman; **assistant manager** Lisa Stanford; **production specialists** Antoinette Hodal, Nichele Johnston, Lori Murphy, Kimberly Oster; **web and new media: senior project manager** Trista Smith, **project leader** Luke Johnson **computer specialists** Walter Jones, Kai Zhang, **web developer** Chris Coleman; **program director, AAAS member central** Peggy Mihelich

director, global collaboration, custom publications, advertising Bill Moran

editor, custom publishing Sean Sanders: 202-326-6430

assistant editor, custom publishing Tianna Hicklin 202-326-6463

associate director, collaboration, custom publications/china/taiwan/korea/singapore Ruolei Wu +86-1367-101-5294

Product (science_advertising@aaas.org); **midwest** Rick Bongiovanni: 330-405-7080, FAX 330-405-7081; **east coast/e. canada** Laurie Faraday: 508-747-9395, FAX 617-507-8189; **west coast/w. canada** Lynne Stickrod: 415-931-9782, FAX 415-520-6940; **uk europe/asia** Roger Goncalves: TEL/FAX +41 43 243 1358; **japan**, Makiko Hara: +81 (0) 3 6802 4616, FAX +81 (0) 3 6802 4615; ads@sciencemag.jp; **china/taiwan** Ruolei Wu: +86 1367 1015 294 rwu@aaas.org

worldwide associate director of *Science* Careers Tracy Holmes: +44 (0) 1223 326525, FAX +44 (0) 1223 326532

Classified (advertise@sciencecareers.org); **u.s./canada/south america** Tina Burks: 202-326-6577; **sales administrator** Marci Gallun; **europe/row sales** Lucy Nelson; **sales assistant** Kelly Grace; **japan** Yuri Kobayashi +81 (0)90-9110-1719; careerads@sciencemag.jp; **china/taiwan** Ruolei Wu: +86 1367 1015 294 rwu@aaas.org; **advertising support manager** Karen Foote: 202-326-6740; **advertising production operations manager** Deborah Tompkins; **senior production specialist/graphic designer** Amy Hardcastle; **production specialist** Yuse Lajiminmuhip; **senior traffic associate** Christine Hall; **sales coordinator** Shirley Young; **marketing manager** Allison Pritchard; **marketing associate** Aimee Aponte

AAAS Board of Directors retiring president, chair William H. Press; **president** Phillip A. Sharp; **president-elect** Gerald R. Fink; **treasurer** David Evans Shaw; **chief executive officer** Alan I. Leshner; **board** Bonnie L. Bassler, May R. Berenbaum, Claire M. Fraser, Elizabeth Loftus, Stephen L. Mayo, Raymond Orbach, Sue V. Rosser, Inder M. Verma

AAAS
ADVANCING SCIENCE, SERVING SOCIETY

SENIOR EDITORIAL BOARD

A. Paul Alivisatos, *Lawrence Berkeley Nat'l. Laboratory*
Cori Bargmann, *The Rockefeller Univ.*
Ernst Fehr, *Univ. of Zurich*
Erin O'Shea, *Harvard Univ.*
Michael S. Turner, *University of Chicago*

BOARD OF REVIEWING EDITORS

Adriano Aguzzi, *Univ. Hospital Zürich*
Takuzo Aida, *Univ. of Tokyo*
Leslie Aiello, *Wenner-Gren Foundation*
Sonia Altizer, *Univ. of Georgia*
Sebastian Amigorena, *Institut Curie*
Angelika Amon, *MIT*
Kathryn Anderson, *Memorial Sloan-Kettering Cancer Center*
Siv G. E. Andersson, *Uppsala Univ.*
Peter Andolfatto, *Princeton Univ.*
Meinrat O. Andreae, *Max Planck Inst., Mainz*
Paola Arlotta, *Harvard Univ.*
Johan Auwerx, *EPFL*
David Awschalom, *Univ. of California Santa Barbara*
Ben Barres, *Stanford Medical School*
Jordi Bascompte, *Estación Biológica de Doñana, CSIC*
Facundo Batista, *London Research Inst.*
Ray H. Baughman, *Univ. of Texas, Dallas*
David Baum, *Univ. of Wisconsin*
Mark Bear, *Massachusetts Inst. of Technology*
Yasmine Belkaid, *NIAID, NIH*
Philip Benfey, *Duke Univ.*
Stephen J. Benkovic, *Penn State Univ.*
Christophe Bernard, *Aix-Marseille Univ.*
Gregory C. Beroza, *Stanford Univ.*
Gabriele Bergers, *Univ. of California, San Francisco*
Peer Bork, *EMBL*
Bernard Bourdon, *Ecole Normale Superieure de Lyon*
Chris Bowler, *Ecole Normale Supérieure*
Ian Boyd, *Univ. of St. Andrews*
Christian Büchel, *Universitätsklinikum Hamburg-Eppendorf*
Joseph A. Burns, *Cornell Univ.*
William P. Butz, *Population Reference Bureau*
Gyorgy Buzsaki, *New York Univ., School of Medicine*
Mats Carlsson, *Univ. of Oslo*
Mildred Cho, *Stanford Univ.*
David Clapham, *Children's Hospital, Boston*
David Clary, *Univ. of Oxford*
Jonathan D. Cohen, *Princeton Univ.*
Robert Cook-Deegan, *Duke Univ.*
James Collins, *Boston Univ.*
Alan Cowman, *Walter & Eliza Hall Inst.*
Robert H. Crabtree, *Yale Univ.*
Wolfgang Cramer, *Mediterranean Inst. of Biodiversity and Ecology*
Jeff L. Dangl, *Univ. of North Carolina*
Tom Daniel, *Univ. of Washington*
Frans de Waal, *Emory Univ.*
Stanislas Dehaene, *Collège de France*
Robert Desimone, *MIT*
Claude Desplan, *New York Univ.*
Ap Dijksterhuis, *Radboud Univ. of Nijmegen*
Dennis Discher, *Univ. of Pennsylvania*
Gerald W. Dorn II, *Washington Univ. School of Medicine*
Jennifer A. Doudna, *Univ. of California, Berkeley*
Julian Downward, *Cancer Research UK*
Bruce Dunn, *Univ. of California, Los Angeles*
Christopher Dye, *WHO*
David Ehrhardt, *Carnegie Inst. of Washington*
Tim Elston, *Univ. of North Carolina at Chapel Hill*
Gerhard Ertl, *Fritz-Haber-Institut, Berlin*
Barry Everitt, *Univ. of Cambridge*
Paul G. Falkowski, *Rutgers Univ.*
Ernst Fehr, *Univ. of Zurich*
Tom Fenchel, *Univ. of Copenhagen*
Michael Feuer, *The George Washington Univ.*
Alain Fischer, *INSERM*
Susan Fiske, *Princeton Univ.*
Anne C. Ferguson-Smith, *Univ. of Cambridge*
Peter Fratzl, *Max Planck Inst.*
Elaine Fuchs, *Rockefeller Univ.*
Wulfram Gerstner, *EPFL Lausanne*
Andrew Gewirth, *Univ. of Illinois*
Karl-Heinz Glassmeier, *TU Braunschweig*
Elizabeth Grove, *Univ. of Chicago*
Kip Guy, *St. Jude's Children's Research Hospital*
Taekjip Ha, *Univ. of Illinois at Urbana-Champaign*
Christian Haass, *Ludwig Maximilians Univ.*
Steven Hahn, *Fred Hutchinson Cancer Research Center*
Gregory J. Hannon, *Cold Spring Harbor Lab.*
Martin Heimann, *Max Planck Inst., Jena*
Yka Helariutta, *Univ. of Finland*
Isaac Held, *NOAA*
James A. Hendler, *Rensselaer Polytechnic Inst.*
Janet G. Hering, *Swiss Fed. Inst. of Aquatic Science & Technology*
Ray Hilborn, *Univ. of Washington*
Michael E. Himmel, *National Renewable Energy Lab.*
Kai-Uwe Hinrichs, *Univ. of Bremen*
Kei Hirose, *Tokyo Inst. of Technology*
David Hodell, *Univ. of Cambridge*
David Holden, *Imperial College*
Lora Hooper, *UT Southwestern Medical Ctr at Dallas*
Jeffrey A. Hubbell, *EPFL Lausanne*
Thomas Hudson, *Ontario Inst. for Cancer Research*
Ray Huey, *Univ. of Washington*
Steven Jacobsen, *Univ. of California, Los Angeles*
Kai Johnsson, *EPFL Lausanne*
Peter Jonas, *Universität Freiburg*
Matt Kaeberlein, *Univ. of Washington*
William Kaelin Jr., *Dana-Farber Cancer Inst.*
Daniel Kahne, *Harvard Univ.*
Daniel Kammen, *Univ. of California, Berkeley*
Joel Kingsolver, *Univ. of North Carolina at Chapel Hill*
Robert Kingston, *Harvard Medical School*
Roberto Kolter, *Harvard Medical School*
Alberto R. Kornblihtt, *Univ. of Buenos Aires*
Leonid Kruglyak, *Princeton Univ.*
Thomas Langer, *Univ. of Cologne*
Mitchell A. Lazar, *Univ. of Pennsylvania*
David Lazer, *Harvard Univ.*
Virginia Lee, *Univ. of Pennsylvania*
Ottoline Leyser, *Cambridge Univ.*
Marcia C. Linn, *Univ. of California, Berkeley*
Jianguo Liu, *Michigan State Univ.*
Luis Liz-Marzan, *CIC biomaGUNE*
Jonathan Losos, *Harvard Univ.*
Ke Lu, *Chinese Acad. of Sciences*
Christian Lüscher, *Univ. of Geneva*
Laura Machesky, *CRUK Beatson Inst. for Cancer Research*
Anne Magurran, *Univ. of St Andrews*
Oscar Marin, *CSIC & Univ. Miguel Hernández*
Charles Marshall, *Univ. of California, Berkeley*
Chris Marshall, *Inst. of Cancer Research*
Martin M. Matzuk, *Baylor College of Medicine*
C. Robertson McClung, *Dartmouth College*
Graham Medley, *Univ. of Warwick*
Yasushi Miyashita, *Univ. of Tokyo*
Richard Morris, *Univ. of Edinburgh*
Edvard Moser, *Norwegian Univ. of Science and Technology*
Sean Munro, *MRC Lab. of Molecular Biology*
Thomas Murray, *The Hastings Center*
Naoto Nagaosa, *Univ. of Tokyo*
James Nelson, *Stanford Univ. School of Med.*
Daniel Neumark, *Univ. of California, Berkeley*
Stuart Newman, *New York Medical College*
Timothy W. Nilsen, *Case Western Reserve Univ.*
Pär Nordlund, *Karolinska Inst.*
Helga Nowotny, *European Research Advisory Board*
Luke O'Neill, *Trinity College, Dublin*
Stuart Newman, *New York Medical College*
N. Phuan Ong, *Princeton Univ.*
Joe Orenstein, *Univ. of California, Berkeley & Lawrence Berkeley National Lab*
Harry Orr, *Univ. of Minnesota*
Andrew Oswald, *Univ. of Warwick*
Steve Palumbi, *Stanford Univ.*
Jane Parker, *Max-Planck Inst. of Plant Breeding Research*
Donald R. Paul, *Univ. of Texas at Austin*
P. David Pearson, *Univ. of California, Berkeley*
John H. J. Petrini, *Memorial Sloan-Kettering Cancer Center*
Simon Phillpot, *Univ. of Florida*
Joshua Plotkin, *Univ. of Pennsylvania*
Philippe Poulin, *CNRS*
Colin Renfrew, *Univ. of Cambridge*
Trevor Robbins, *Univ. of Cambridge*
Jim Roberts, *Fred Hutchinson Cancer Research Ctr.*
Barbara A. Romanowicz, *Univ. of California, Berkeley*
Jens Rostrup-Nielsen, *Haldor Topsoe*
Mike Ryan, *Univ. of Texas, Austin*
Shimon Sakaguchi, *Kyoto Univ.*
Miquel Salmeron, *Lawrence Berkeley National Lab*
Jürgen Sandkühler, *Medical Univ. of Vienna*
Alexander Schier, *Harvard Univ.*
Randy Seeley, *Univ. of Cincinnati*
Vladimir Shalaev, *Purdue Univ.*
Joseph Silk, *Institut d'Astrophysique de Paris*
Denis Simon, *Arizona State Univ.*
Alison Smith, *John Innes Centre*
Davor Solter, *Inst. of Medical Biology, Singapore*
Peter Sorger, *Harvard Medical School*
John Speakman, *Univ. of Aberdeen*
Allan C. Spradling, *Carnegie Institution of Washington*
Jonathan Sprent, *Garvan Inst. of Medical Research*
Paula Stephan, *Georgia State Univ. and National Bureau of Economic Research*
Elsbeth Stern, *ETH Zürich*
V. S. Subrahmanian, *Univ. of Maryland*
Ira Tabas, *Columbia Univ.*
Yoshiko Takahashi, *Kyoto University*
Sarah Teichmann, *Cambridge Univ.*
John Thomas, *Duke Univ.*
Herbert Virgin, *Washington Univ.*
Bert Vogelstein, *Johns Hopkins Univ.*
Cynthia Volkert, *Univ. of Göttingen*
Bruce D. Walker, *Harvard Medical School*
Douglas Wallace, *Dalhousie Univ.*
Ian Walmsley, *Univ. of Oxford*
David A. Wardle, *Swedish Univ. of Agric Sciences*
David Waxman, *Fudan Univ.*
Jonathan Weissman, *Univ. of California, San Francisco*
Kathy Willis, *Oxford Univ.*
Ian A. Wilson, *The Scripps Res. Inst.*
Timothy D. Wilson, *Univ. of Virginia*
Rosemary Wyse, *Johns Hopkins Univ.*
Jan Zaanen, *Leiden Univ.*
Kenneth Zaret, *Univ. of Penn. School of Medicine*
Jonathan Zehr, *Univ. of California, Santa Cruz*
Maria Zuber, *MIT*

BOOK REVIEW BOARD

John Aldrich, *Duke Univ.*
David Bloom, *Harvard Univ.*
Angela Creager, *Princeton Univ.*
Richard Shweder, *Univ. of Chicago*
Ed Wasserman, *DuPont*
Lewis Wolpert, *Univ. College London*

The Human Diseasome

Albert-László Barabási,
Distinguished Professor of Physics

Meet **Albert-László Barabási.** He and his world-leading network science team are revolutionizing medicine by mapping the molecular and genetic links between diseases—the human "diseasome." It's a breakthrough that will give physicians and researchers alike a powerful new tool to forecast and treat disease.

Learn how Barabási and our other faculty researchers are making the world healthier. Northeastern University: making tomorrow happen.

northeastern.edu/tomorrow

Northeastern University **Making Tomorrow Happen**

AROUND THE WORLD

Washington, D.C. **1**

House Passes Helium Bill

On 26 April, the U.S. House of Representatives passed a bill by a vote of 394 to 1 that would head off a looming helium shortage.

Helium is required to run MRI machines, manufacture optical fibers and microchips, and cool samples to near absolute zero. Federal helium reserves now supply 42% of the United States' demand for helium and 35% of the global demand.

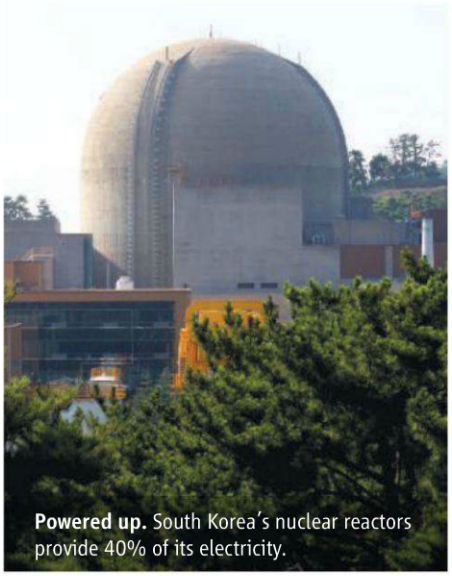
The shortage is of the government's own making. Following a mandate by Congress, in 2003 the Bureau of Land Management (BLM) began to sell off federal helium reserves stored near Amarillo, Texas, to recoup the $1.3 billion that the government spent accumulating the helium. But by October, BLM will break even and has no further mandate to sell the remaining helium. BLM also charges a below-market price for its helium, which encourages waste and discourages the development of new sources of helium.

The House bill would continue sales for another year and then would require at least 60% of the helium to be sold in semiannual auctions. After the reserve drops below a certain level, in about 5 years, sales would be limited to federal users. The Senate Committee on Energy and Natural Resources will hold a hearing on a similar bill on 7 May. **http://scim.ag/Hebill**

Washington, D.C., and Seoul **2**

Extension of Nuclear Deal Disappoints Scientists

The announcement on 24 April that U.S. and South Korean negotiators seek to extend the conditions of a long-term agreement on nuclear power cooperation is not pleasing anyone. Under the 1974 deal, due to expire next year, the United States provides expertise and fuel to South Korea's nuclear power industry as long as the latter refrains from enriching uranium and reprocessing spent fuel. But South Korea wants to enrich uranium to supplement its growing nuclear power plant export business, and it wants reprocessing technology to handle spent fuel from its 23 reactors, which produce nearly 40% of its electricity. Such know-how, however, can be applied to producing nuclear bomb material.



Powered up. South Korea's nuclear reactors provide 40% of its electricity.

Unable to reach terms on a long-term deal, U.S. and Korean negotiators opted to extend the current agreement for 2 years. That doesn't please scientists. "Frankly, the nuclear power research community is disappointed," writes nuclear fusion specialist Gyung-Su Lee to *Science* in an e-mail. Lee, who led the development of South Korea's Korean Superconducting Tokamak Advanced Research reactor, noted that South Korea has adhered to the Nuclear Non-Proliferation Treaty and should be considered as trustworthy as Japan, which is developing enrichment and reprocessing technologies. **http://scim.ag/USKornuc**

Bethesda, Maryland **3**

HIV Vaccine Study Meets Premature End

The largest HIV vaccine study in the world came to a premature halt last week when an early analysis showed that it didn't work. The trial, launched in 2009, was taking place in 19 U.S. cities and involved giving two vaccines to 2500 men and transgender women who have sex with men.

The vaccines, made by the U.S. National Institute of Allergy and Infectious Diseases (NIAID), consisted of HIV genes stitched into DNA and, separately, a harmless adenovirus. NIAID halted the trial 2 years before its planned end upon the recommendation of an independent data and safety monitoring board. The board found that vaccinated people became infected as readily as those given placebo shots. The vaccines also failed to help control the virus in those who became infected.

Mitchell Warren, head of the AIDS Vaccine Advocacy Coalition in New York City, said the silver lining is that the study's early end tells the field to look elsewhere immediately. "There are other approaches that must be pursued without delay," Warren said.

Ontario, Canada **4**

Experimental Lakes Get Lifeline

Canada's Experimental Lakes Area (ELA) will remain open this year with money from the province of Ontario. But that funding, announced on 24 April by Ontario Premier Kathleen Wynne, may be only a temporary lifeline. The facility has conducted ecosystems research at 58 lakes since 1968. But last year, the Canadian government axed its $2 million annual appropriation, gave notice that it planned to start tearing down structures in September, and hired security guards to keep scientists off the site after 31 March.

Wynne's announcement gives ELA a reprieve, but the provincial government has not decided how much money it will free up for operating costs, says Laurel Broten, Ontario's minister of intergovernmen-

CREDITS (TOP TO BOTTOM): © THEODORE GRAY/VISUALS UNLIMITED/CORBIS; YONHAP NEWS/YNA/NEWSCOM

tal affairs. Roughly $600,000 is needed to operate core ELA facilities annually, while an additional $1.4 million is needed to cover the cost of salaries for ELA staff members that carry out scientific projects or conduct the routine work of taking water samples and monitoring water flows.

Quick save. ELA reseachers get a reprieve.

Negotiations are ongoing with the federal government and the province of Manitoba "to find the best arrangement." It's hoped that ELA can ultimately survive under the umbrella management of the International Institute for Sustainable Development. **http://scim.ag/ELAOnt**

Brussels **5**

Controversial Pesticide Ban Goes Forward

The European Commission is going ahead with a 2-year moratorium on three widely used pesticides that are potentially harmful to bees, although E.U. countries remain split on the issue. The decision follows reports published in January by the European Food Safety Authority (EFSA) concluding that the three neonicotinoids pose an "acute risk" to honey bees that are essential to farming and natural ecosystems.

After two inconclusive votes by member states, the commission is allowed to drive its proposal through. Those votes occurred on 15 March, in the Standing Committee on the Food Chain and Animal Health, and in an appeals committee on 29 April. Now, the commission has decided to use its right to go ahead with the plan.

"The Commission's role is to protect consumers. EFSA's reports have revealed a risk, so we're taking time to step back and assess the situation," says a Brussels source close to the negotiations. The 2-year restriction to pesticide use will apply to plants attractive to bees (maize, cotton, sunflower, and rapeseed) starting 1 December. It will apply across the European Union, even in the countries, including the United Kingdom, that opposed the moratorium. **http://scim.ag/ECbees**

Priština **6**

Reconciliation With Serbia Opens Research Doors

Three days after leaders from Serbia and Kosovo reached an E.U.-brokered agreement on 19 April, the European Commission formally proposed to let Kosovo participate in 22 E.U. programs, including the €55 billion Seventh Framework Programme (FP7) for R&D, Europe's satellite navigation program Galileo, and the European Earth-monitoring system GMES. FP7 ends this year, but the agreement is expected to also apply to its successor, Horizon 2020, which covers the next 7 years.

If E.U. member states agree, Kosovo will move from taking part in research programs as a "third country" to becoming an "associated country." In return for paying a fee to the European Union, its representatives can take part in program management committees, and its organizations and proposals will receive the same treatment as those from E.U. member states.

The move would give the small Balkan country fresh opportunities to shore up its minuscule research effort—but it may have to invest more to benefit from them. According to the country's own National Research Programme, published in 2010, R&D remains a "marginal undertaking" in Kosovo, with "the absence of any critical mass of research and technological development … funding for at least the last 20 years." **http://scim.ag/KosovoEU**

>>

Obama Wishes Happy 150th to U.S. National Academy of Sciences

President Barack Obama had an appreciative audience at the U.S. National Academy of Sciences (NAS) on 29 April when he delivered a speech celebrating the group's 15 decades of service as the government's technical adviser. "For 150 years, you've strived to answer big questions, solve tough problems, not for yourselves but for the benefit of the nation," he told a crowd of science luminaries.

Obama also took a swipe at congressional efforts to impose controversial new grant-making criteria on U.S. science agencies (see p. 534). "I will keep working to make sure that our scientific research does not fall victim to political maneuvers or agendas that in some ways would impact on the integrity of the scientific process," he promised. And he drew a laugh by noting that academy members work for free. "[P]art of what's made the Academy so effective is that all the scientists elected to your elite ranks are volunteers," he said. "Which is fortunate because we have no money anyway."

CREDITS (TOP TO BOTTOM): CRADEL/WIKIMEDIA COMMONS; ELA; RACHEL BRODY/NATIONAL ACADEMY OF SCIENCES

Will 'Volcanic Winter' Debate Now Cool Down?

About 75,000 years ago, Mount Toba in Indonesia blew its top in the biggest volcanic eruption of the past 2 million years. The explosion spread ash across Asia, possibly blocking sunlight and cooling the continent for up to a decade. Some researchers have suggested that ash from Toba also blew over East Africa, driving early *Homo sapiens* nearly to extinction. This "volcanic winter" scenario is hotly debated, however (*Science*, 5 March 2010, p. 1187).

This week, in the *Proceedings of the National Academy of Sciences*, a team of researchers from the University of Oxford in the United Kingdom and the University of Minnesota report finding traces of ash from the Toba explosion in cores drilled into the floor of Lake Malawi, 7000 kilometers west of Toba—the farthest from the volcano yet recorded. However, these microscopic fragments of glass—sourced to the Toba explosion by their chemical composition—seem to have posed little threat to our ancestors. The composition of algae and other organic matter in sediments above and below the ash layer give no hints of dramatic temperature drops or other environmental changes at that time, the team concludes.

>> AROUND THE WORLD

London 7

New Libel Law for England And Wales

A defamation bill that advocates say will help protect free speech—including statements in peer-reviewed scientific publications—received the go-ahead from Parliament on 24 April. Under current English law, plaintiffs alleging spoken defamation or published libel need only show that a public statement might inflict reputational damage. Under the new law, claimants will have to present evidence of "serious financial harm."

Reform efforts were prompted by several recent cases—including that of Simon Singh, a science writer unsuccessfully sued by the British Chiropractic Association in 2008 over a column that he wrote for *The Guardian* questioning the effectiveness of chiropractic treatments, and another 2008 case in which the journal *Nature* was dragged into a costly court battle by an Egyptian researcher who claimed his reputation was damaged by a *Nature* article that alleged that he had self-published many papers without peer review.

The government still has to clarify how the law will work in practice, which will help determine the cost of future libel cases. Another area that awaits clarification is whether the new protections will extend to material published on the Internet, Singh says—which would "turn a good bill into a great bill." **http://scim.ag/UKlibel**

Science LIVE

Join us on Thursday, 9 May, at 3 p.m. EDT for a live chat on **new exoplanet discoveries**. How close are we to finding a mirror Earth? **http://scim.ag/science-live**

NEWSMAKERS

Tropical Medicine Researcher To Lead Wellcome Trust

Farrar

Infectious diseases researcher **Jeremy Farrar** will take over the reins at the Wellcome Trust, the United Kingdom's most important private funder of biomedical research, on 1 October. Endowed with more than $22 billion, the Wellcome Trust spends about $1 billion annually.

Farrar is now heading the Oxford University Clinical Research Unit in Ho Chi Minh City, Vietnam, where he has done research on drug resistance in tuberculosis and other diseases. He is "an extraordinary guy," says Nicholas White, a malaria researcher at the University of Oxford in the United Kingdom and Mahidol University in Bangkok. "He understands science and he understands people."

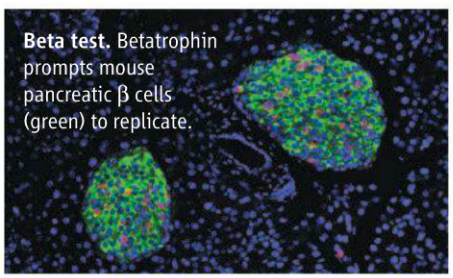
In January 2004, Farrar was part of a team that diagnosed the first human case of H5N1 avian influenza in Vietnam. In recent weeks, he has collaborated in an effort to analyze travel patterns between China and the rest of the world, which may help predict how the novel H7N9 influenza virus might spread.

Farrar succeeds Mark Walport, who had left the post in March to become the U.K. government's chief science adviser. Ted Bianco, the trust's director of technology transfer, will continue to serve as acting director until October. **http://scim.ag/Farrar**

FINDINGS

New Liver Hormone To Treat Diabetes?

A hormone in mice prompts the rodents to boost their production of pancreatic β cells—the ones that make insulin and



Beta test. Betatrophin prompts mouse pancreatic β cells (green) to replicate.

CREDITS (TOP TO BOTTOM): © CHARLES & JOSETTE LENARS/CORBIS; WELLCOME LIBRARY, LONDON, WELLCOME IMAGES; PENG YI

are missing or not productive enough in patients with diabetes. The hormone, called betatrophin, is made by people as well; if the effect in the human pancreas is similar, betatrophin could be a new therapy for diabetes, the researchers say.

Insulin helps cells take up glucose, a key cellular energy source. Many patients with diabetes inject themselves with insulin to make up for what the pancreas can't produce.

While studying how a healthy pancreas produces β cells, cell biologists Douglas Melton, Peng Yi, and Ji-Sun Park at Harvard University identified a gene activated in the liver and fat tissues of mice. The gene codes for a protein excreted by cells, suggesting that it might be a hormone. They called it betatrophin.

When the researchers injected copies of the betatrophin gene into the livers of normal mice, the animals' pancreases made as much as 30 times more β cells than usual, the team reported online on 25 April in *Cell*. After a week, the injected mice had more than twice the number of β cells as animals that didn't get the extra genes.
http://scim.ag/livdiab

Bird queue. A chicken gets a shot of H5N1 vaccine.

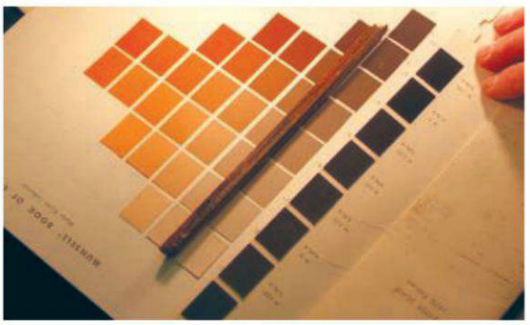
Gene Swap Helps Bird Flu Spread Between Mammals

A simple gene swap can make the dangerous bird flu strain H5N1 transmissible between mammals, finds a paper published by *Science* this week.

H5N1 can infect humans in contact with infected birds, but does not spread efficiently between people. A team led by Chen Hualan at the Harbin Veterinary Research Institute in China wanted to know if that might change if H5N1 picks up genes from H1N1, a human flu strain, when the two infect the same host. They created 127 hybrids of the two viruses and tested which would spread from one guinea pig to another via respiratory droplets. H5N1 needed to borrow only one gene from H1N1, they found, to efficiently spread between the guinea pigs. A similar study published last month in *PLOS ONE* by Ron Fouchier's group at Erasmus MC in Rotterdam, the Netherlands, found that these gene swaps alone didn't make the virus airborne between ferrets, another animal model for flu.

Virologist Simon Wain-Hobson of the Pasteur Institute in Paris says the study is a "super piece of work" scientifically, but also calls it "very dangerous" because the virus could escape from the lab, and the paper could help wannabe bioterrorists.
http://scim.ag/geneswap

Random Sample



Funeral Colors

On 21 April 1865, 7 days after the assassination of President Abraham Lincoln, his funeral procession departed Washington, D.C., aboard a nine-car train bound for Springfield. Millions of mourners turned out to see the train pass by and pay their respects in what has been called the greatest funeral in the history of the United States.

Despite all the eyewitness accounts, one small detail was lost to history: the precise color of the funeral car. Various accounts referred to it as "rich chocolate brown" and "claret red." The car itself was all but destroyed in a fire in 1911. Now, from the few fragments that remain, researchers from the University of Arizona say that they've solved the color mystery.

Chemist Wayne Wesolowski had previously reconstructed a model-sized version of the railcar (top) as director of the 10-year Lincoln Train Project at Benedictine University in Chicago, Illinois. His "best guess" at the color for the model, he says, was based on the "rich chocolate brown" description. "In 1865, there was no milk chocolate—chocolate was a drink you got in coffee houses, more like Dutch chocolate. It was a rich, dark maroon color."

Now, Wesolowski says, a group known as the Lincoln Funeral Car Project is constructing a full-scale replica of the train, which will reproduce the route of the 16th president's final journey for the upcoming 150th anniversary. To determine the right color for the car, Wesolowski and his team compared microscopic paint fragments from a surviving pencil-sized fragment of the car to national paint standards (bottom). The original color, they found, was a somber maroon—just slightly darker than Wesolowski used for his model.

CREDITS (TOP TO BOTTOM): W. WESOLOWSKI (2); SI CHUAN/EPA/NEWSCOM

Under scrutiny. Representative Lamar Smith wants to change peer review at NSF.

U.S. SCIENCE POLICY

Bill Would Set New Rules For Choosing NSF Grants

The new chair of the House of Representatives science committee has drafted a bill that, in effect, would alter peer review at the National Science Foundation (NSF) by embracing a set of criteria chosen by Congress. The proposed legislation, from Representative Lamar Smith (R–TX), has angered congressional Democrats and left NSF officials wondering what Smith thinks is wrong with the current system.

The bill, called the "High Quality Research Act," is the latest—and bluntest—attack by congressional Republicans on how NSF decides which grant proposals to fund. In late March, Senator Tom Coburn (R–OK) successfully attached language to a 2013 spending bill that prohibits NSF from funding any political science research for the rest of the fiscal year unless its director certifies that the research pertains to economic development or national security.

Smith's draft bill, which has not yet been formally introduced, goes much further by applying similar language to NSF's entire research portfolio. In particular, it would require the NSF director to certify, prior to any award, that the work being funded is "ground breaking" and "not duplicative" as well as important to national interests. NSF's current guidelines, updated last year, ask reviewers to consider the "intellectual merit" of a proposed research project as well as its "broader impacts" on the scientific community and society.

"I don't know what problem he's trying to solve," says Dan Arvizu, chair of the National Science Board, a 24-member presidentially appointed body that sets policy for the foundation. "NSF's mission is to support the best research, and we feel it has done a good job doing that. There's always room for improvement, but the legislation implies that you're being unpatriotic if you don't agree to do things differently."

Arvizu notes that the situation "seems to have escalated" since Republicans on the science committee grilled acting NSF Director Cora Marrett and John Holdren, the president's science adviser, during back-to-back hearings on 17 April. During those hearings, which examined the Obama administration's proposed 2014 budget for NSF and other science agencies, members read the titles of several NSF grants in the social and political sciences and asked both administration officials to explain why the research was funded.

Last week, Smith asked Marrett for more information on five of those recent grants. In particular, he requested copies of the comments from each reviewer, as well as the notes of the NSF program officer managing the awards.

"I have concerns regarding some grants approved by the Foundation and how closely they adhere to NSF's 'intellectual merit' guideline," Smith wrote in his letter. In a subsequent statement, Smith said that the designated grants "do not seem to meet the high standards of most NSF-funded projects. Congress has a responsibility to review questionable research paid for by hard-working American taxpayers. If academic or other institutions want to conduct such research on these kinds of subjects they can pay for them with their own private funds."

Smith's letter didn't sit well with the top Democrat on the science committee, Representative Eddie Bernice Johnson (D–TX). The next day she sent him a blistering missive, questioning his judgment and his motives and asking him to withdraw his request.

"In the history of this Committee, no Chairman has ever put themselves forward as an expert in the science that underlies specific grant proposals funded by NSF," Johnson wrote. "I have never seen a Chairman decide to go after specific grants simply because the Chairman does not believe them to be of high value."

Coincidentally, President Barack Obama spoke earlier this week at the annual meeting of the National Academy of Sciences, which is celebrating its 150th anniversary. Obama, who also addressed NAS members shortly after taking office in 2009, praised them for their "passion" and commitment to improving people's lives. Although he didn't mention Smith or the pending legislation, Obama promised to protect "our rigorous peer review system" and asked the audience to help him ensure that federally funded research "does not fall victim to political maneuvers or agendas that in some ways would impact on the integrity of the scientific process."

After the speech, Representative Rush Holt (D–NJ) suggested that scientists and politicians can wind up at cross-purposes even if they agree that research is important. "Members of Congress, like the people they represent, value the fruits of science," says Holt, a former plasma physicist. "But they have hardly a clue of how the research process works and what is needed to sustain it."

Arvizu says that he and Marrett have offered to meet with Smith and Representative Larry Bucshon (R–IN), who chairs the committee's research panel, "so we can understand the intent of the proposed legislation. There is nothing more important on my calendar," Arvizu says, "than to clear up any confusion they might have about how NSF conducts peer review." Arvizu says that there "are a number of things wrong with what [Smith] has proposed" but that he wanted to avoid "a public debate over the criteria."

–JEFFREY MERVIS

CREDIT: BILL CLARK/CQ ROLL CALL PHOTOS/NEWSCOM

PRIMATE STUDIES

Research in Limbo as Harvard Moves to Close Center

Harvard Medical School abruptly announced last week that it is closing a major primate research center that has been funded by the National Institutes of Health (NIH) for 50 years. Although the center had been under investigation for its handling of animals, Harvard said that it is closing the center within 2 years because the outlook for federal research funding is bleak. Many in the small community of researchers who work with nonhuman primates were dismayed, noting that their colleagues will have to find new homes for their projects, slowing work on AIDS and other diseases.

"This is a tragedy. An AIDS vaccine or cure is not going to happen without nonhuman primate research. We need more of this, not less," says Louis Picker of Oregon Health & Science University in Portland.

Harvard leaders said that they decided not to apply to renew the New England Primate Research Center's (NEPRC's) 5-year, $10 million-a-year base grant from NIH because tight NIH funding has forced the medical school to make strategic reductions. But some current and former NEPRC researchers, who say that they were given no warning, put the blame on medical school and university decision-makers. Critics say that a heavy-handed approach to resolving animal care problems spurred a "death spiral" of faculty departures and put the center on precarious financial footing. About a dozen of the center's faculty members have left or accepted outside offers in the past 2 years.

"I'm furious," says David Watkins, an immunologist at the University of Miami in Florida and a close colleague of NEPRC researcher Ronald Desrosiers, who stepped down as director in 2011. "This was a self-inflicted wound by the Harvard administration, and it is disingenuous to suggest that the primate center is closing because of reduced NIH funding," Watkins says. (Desrosiers, who will join the University of Miami in June, declined to speak to *Science* on the record.) "I think it's sad, and I strongly suspect that there's more to the decision than finances," says Welkin Johnson, who left NEPRC for Boston College in 2011 for reasons unrelated to the recent problems.

NEPRC, located in Southborough, Massachusetts, 25 miles west of Boston, is one of eight NIH-funded national primate centers. About 20 faculty members, 32 graduate students and postdocs, and dozens of outside researchers from Harvard and other institutions use its nearly 2000 rhesus macaque monkeys and other primates for studies including infectious diseases, drug addiction, and primate behavior. NEPRC researchers discovered the simian form of the HIV virus, SIV, in 1985 and later did the first work showing that an AIDS vaccine could be protective. Five years ago, NEPRC's proposal to renew its base grant received the highest score ever for a primate center, researchers there say.

Then, in June 2010, a cotton-top tamarin was found dead in a cage that had gone through a washer. Although an autopsy showed that the animal had died of natural causes, the incident sparked an investigation that found problems in compliance with animal care rules, such as unapproved blood draws. In September 2011, the medical school responded with severe measures: An administrator, two chief veterinarians, and Desrosiers were replaced.

Lower priority. Costs—not animal care or faculty issues—drove a decision to close the New England Primate Research Center, Harvard says.

CREDIT: *THE BOSTON GLOBE*/GETTY IMAGES

The problems only escalated, however. In the next 6 months, three more monkeys died, two because they lacked water. After a U.S. Department of Agriculture investigation criticized the center again in February 2012, interim director Frederick Wang stepped down and Harvard assigned a blue ribbon panel to review NEPRC's animal care procedures. "They had some pretty significant problems," says panel member Andrew Lackner, director of an NIH primate center at Tulane University in Covington, Louisiana. When the panel's report came out in August 2012, NEPRC had begun putting in place new training and operating procedures. "Most everyone thought they had turned the corner," Lackner says.

By that time, however, faculty members had begun to leave. Several former or departing researchers fault Harvard for not reversing the exodus. Harvard Medical School Dean Jeffrey Flier says that it would not have made sense to recruit "at a time when we were fixing the place and making its operations and leadership what we felt it had to be." Harvard launched a search for a new director last year and chose an internal candidate, NEPRC interim director R. Paul Johnson.

Despite NEPRC's problems, Flier says that Harvard was ready to rebuild a few months ago. "The plan was to make a substantial investment," he says. Flier expected that the center's NIH base grant would be renewed. But after reviewing priorities in the light of the U.S. budget pinch, Harvard instead made the "very difficult decision" to close NEPRC, according to Flier: "We don't feel that a primate center was required for the ongoing activities of Harvard Medical School." Fliers denies that animal activists influenced the decision. Some onlookers suggest that because the medical school is in debt, a rescue would have required tens of millions of dollars from Harvard University, which may not have been as supportive of the troubled center.

Harvard expects to move the center's animals to other primate centers and sites over the next 1 to 2 years; no animals will be sacrificed. James Anderson, an NIH deputy director, says of the 130 or so research projects at the center: "This is important science, and we will work with the other [NIH] centers to transfer what we can transfer." Asked whether this will lead to restructuring or trimming the $87 million budget for primate centers, Anderson said: "It's too early to tell."

–JOCELYN KAISER

NEWSMAKER INTERVIEW

Bill and Melinda Gates Talk Science

The philanthropists Bill and Melinda Gates were in Washington, D.C., last weekend talking about the power of science to save and improve the lives of the poor. The occasion was the 150th annual meeting of the National Academy of Sciences, where they received the academy's most prestigious honor, the Public Welfare Medal, shared by the likes of Norman Borlaug, Donald A. Henderson, William Foege, and John D. Rockefeller Jr. The Gateses were commended for changing the trajectory of international public health through their foundation, the Bill & Melinda Gates Foundation, which since it began in 2000 has awarded more than $20 billion in grants in global health and development.

Before picking up their award, Bill and Melinda sat down for a conversation with *Science* in the academy's august building on Constitution Avenue Northwest. They touched on many topics, including why vaccines are the best buy in public health, the urgent need to boost routine childhood immunization in places like northern Nigeria, where coverage is less than 30%, the importance of working with local communities, and especially women, to be sure they get culturally appropriate technologies they want. Mostly they talked about two of their passions: eradicating polio worldwide by 2018, and getting modern contraceptives to another 120 million women by 2020. The Gateses also talked about how the foundation decides what to fund and the most exciting research ideas that they have heard in the past few years.

This interview had been edited for brevity and clarity. The full transcript is available online at http://scim.ag/GatesQA.

–LESLIE ROBERTS

Q: You just came back from Abu Dhabi where you committed another $1.8 billion for a $5.5 billion plan to eradicate polio over the next 6 years. My question is, why is it worth it?

B.G.: With polio, you have two ways you can go. You can either give up and let the disease spread back. So you'd have 100,000 or more kids being crippled if you didn't have special money for this. And you'd still have to be giving polio vaccines or you'd get even bigger numbers.

Or you can go forward and finish the thing, in which case then you no longer have to do any of the vaccination campaigns or even routine [immunization] for polio. So you save tens of billions of dollars—it's an incredible deal. And that money and effort will get pushed over to other global health activities.

Q: How much of success in polio eradication is going to depend on the science and technology and how much on the boots-on-the-ground work of getting the polio vaccine to all the children?

B.G.: There's quite a bit of science in this. We've gone and taken all of the campaign activities and created the digital database, which you might be amazed to hear wasn't there before. We've really looked at the force of infection, this very rigid disease modeling thing that the foundation funds. Then, we take the genetics [of the virus], we sequence some or all of every one of these cases. You can do inheritance trees to show where the virus is coming from.

In Nigeria, we use satellite photos to find villages; we use GPS trackers put in the vaccine boxes to see where the [vaccination] teams are going. At the end of the day, you can compare where the teams went to the microplan [spelling out where they were supposed to go].

M.G.: This is huge compared to just a few years ago. We thought vaccinators were going around to specific huts and houses and villages and it just wasn't happening.

And you can't underestimate the momentum coming out of India—the fact that they actually got it done. That allows us to go to Nigeria and be really serious about what we learned from India and what can we replicate here—all of the issues that you're talking about from science to technology to human resources. We're not naive now, as a community, about what all of those pieces take and how incredibly difficult it is to do. But the fact the countries know it is possible now helps a lot.

B.G.: Overall, the polio [case] numbers are actually pretty good globally this year. It's still not going to be easy, but certainly the foundation has put its credibility behind this. But the challenges of people who mistakenly attack [the program] are also greater than they've ever been. The amount of bad rumors and the violence.

High impact. On 28 April, the National Academy of Sciences honored Bill and Melinda Gates for using science to improve the lives of millions of the world's neediest people.

Q: The security challenges are greater?

B.G.: Greater in terms of the people going after it. We really haven't had much in the way of violence against polio workers until the last 6 months. And we've had it in both Pakistan and Nigeria.

India was harder [biologically]. The birth cohort is 27 million. Nigeria is 6 million. Pakistan is 5 million. But the one thing we never had [in India] was the violence. We had refusal; we had misunderstanding; we had fatigue. But we didn't have the violence.

Q: The global polio eradication effort now costs $1 billion a year. If it is not done in 2018, or 2019, or 2020, what happens? Is there a time when you pull the plug?

M.G.: We've had to make this decision several times now. The amount for the foundation is going up, but it's our belief that it can get done. Bill and I will sit down every 2 years to decide, is that the investment we want to make? So our next decision point on that luckily is a lot further out, it isn't until 2016. Hopefully [by then] we'll have made a lot of progress, and we won't have to make that decision.

Q: I understand that contraception and access to family planning are one of your top priorities.

CREDIT: AP PHOTO/JACQUELYN MARTIN

M.G.: Definitely. The thing that struck me is, as I would go around and talk to women about vaccines for their children, if you really sit and stay with them long enough, they will eventually bring the conversation around and ask, "Why can't I get my shot to keep me from having children?" And they were extremely vociferous about it. Women will tell you they need to be covert. Their husband thinks they've gone to the field for the day, which is their normal activity, and then they'll divert and go to the clinic.

The predominant thing that women use in Africa is Depo-Provera, the shot they get every 3 months. The stock outs are hugely problematic. And to me, it's also a question of science, which is why don't we make the shot so that it lasts a whole lot longer? Well, that has to do with the fact that we don't do research around these hormonal contraceptives anymore.

Women are starting to use more implants in their arms in Africa, and we have to make them more widely available. They're becoming a bit more acceptable for two reasons. One is the rods are much, much smaller than they were 5 years ago and you only need two. So getting them in is easier when they go to the health clinic. And unless their husband knows to feel around for them, they're not obvious in their arms. Women will tell you that if their husband is smart he will feel around for it. I've literally had women tell me I'm willing to be beaten up over it though because by then it's in my arm and it lasts 5 years.

Q: How much is the foundation spending on family planning?

M.G.: Our recent commitment at the London Summit on Family Planning is over $560 million, and between now and 2020 it adds up to almost $1 billion. The goal of the summit is to get another 120 million women voluntary access to family planning by 2020. It's a stretch goal, but we think it's doable by 2020.

Q: I'm curious about how the foundation sets priorities.

M.G.: I call it an economic approach. We look at what the biggest adult killers are in the developing world, and the biggest childhood killers. At the start of the foundation we asked which of these diseases—or which of these issues if it's children—can we affect for the fewest dollars? And that's how we came to vaccines first and foremost.

People think that we're very science-focused, which we absolutely are. The other thing, though, is it's not just science—the biology piece of science.

Seven million children still die a year, and 40% die in the first 30 days of life. Vaccines don't help you in the first 30 days of life. The three things that help in the first 30 days of life are kangaroo care—that is, keeping the baby warm immediately after it's born—immediately and exclusively breastfeeding the baby; and clean and proper cord care. That's all the social sciences.

First we worked with a researcher at Johns Hopkins [Center for Global Health] named Vishwajeet Kumar, who lives and works in Uttar Pradesh. He was able to prove that in these villages in northern India those three things bring down death by over 50%. So once we knew that was true, then our work became, how do you spread these things through local grassroots NGOs? And how do you eventually get it so that the women are actually spreading those best practices? That's how we're going to get at that first 30 days of life, not with a hardcore technology tool.

Q: How involved are you personally in evaluating research proposals?

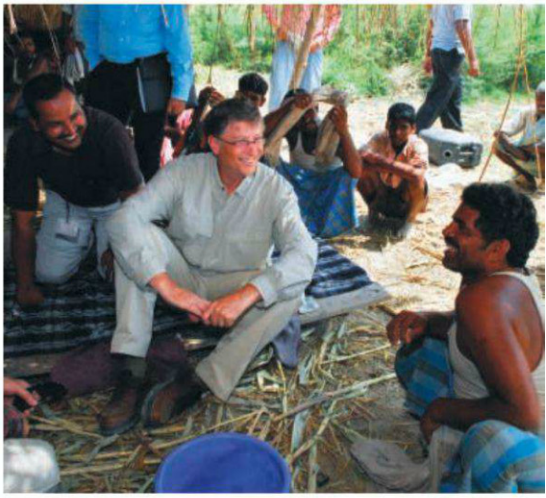
B.G.: Well, we have a malaria team, a diarrhea team, a pneumo team, mother and child team, nutrition team, a [family] planning team, and a lot of the grants are reviewed by them. There'll be some that we talk about, like which malaria vaccine constructs do we [use], and what are we doing with the AIDS vaccine constructs? There are some unusual things like mosquito genetics we look at, because that's kind of a novel thing. We meet with a lot of scientists.

Q: How does an individual scientist get your attention and get your money?

B.G.: We have these grand challenges. You send us a 2-page proposal.There's a pretty small amount of money at the first level—it's $100,000. Then if that works, then it's a million. And it goes up from there, although the drop-off rate, it's pretty narrow by the time you get to the large amounts. We've put out grand challenges on toilets, malaria, contraception. So we're nothing like the NIH [National Institutes of Health]. We're funding in high-impact, poor countries, disproportionate disease burden activities.

Q: What are the most exciting ideas in global health research that you've heard about in the past year?

B.G.: There are a lot of new approaches to vaccines, like DNA and RNA vaccines. Often we have trouble with [vaccine] delivery, so if you can have a depot [in the body] where you'd only have to give the vaccine once and then it would release over time instead of having to get the child vaccinated three times. Or even for drugs—instead of having to take AIDS or TB medicine daily, you had some drug depot that over a period of 45 days it would release it.



Fieldwork. Bill Gates in Bihar state in India (*above*); Melinda Gates (kneeling, *below*) in Dangbo, Benin.

M.G.: I think the science around the sanitation stuff is pretty cool myself.

B.G.: The reinventing the toilet?

M.G.: Yes. I mean the fact that we finally have scientists—hardcore scientists—thinking about what you do with that waste, and do you convert it into other forms of energy, and could you really do that for the developing world? If we wait for the flush toilet to get around to the developing world, that is a long time and it's very expensive.

So coming up with something that doesn't smell, that the waste could actually be reusable so it's useful to the community, and that we have great scientists' minds around that, I might not have thought 5 years ago that was so exciting, but I think it's really exciting.

CREDITS (TOP TO BOTTOM): AP PHOTO/PRASHANT RAVI; AP PHOTO/ERIC FEFERBERG

eppendorf

New accessory:
Eppendorf
ThermoTop®

Multiple Talents

Eppendorf ThermoMixer™ and Eppendorf ThermoStat™

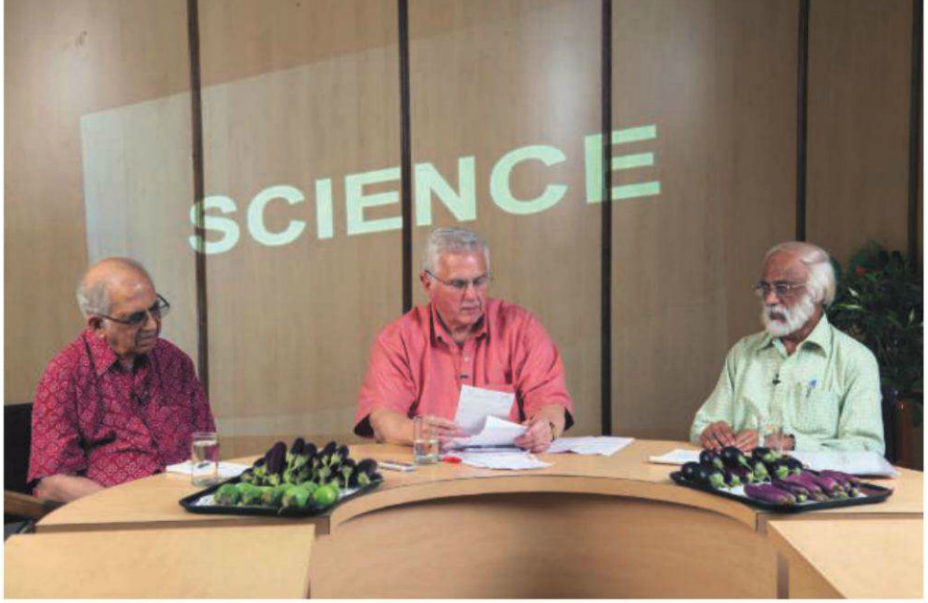
The new generation of Eppendorf Temperature Control and Mixing Instruments offer you more than what meets the eye.

Not only do they provide the versatility and accurate temperature control you come to expect, they now also offer you superior mixing performance, outstanding ergonomic operation and a unique ThermoTop.

> ^{2D}Mix-Control: perfect mixing results in all vessel formats
> Eppendorf QuickRelease™: ergonomic block exchange in just 2 seconds
> Intuitive operation: pre-optimized program and temperature keys
> Eppendorf ThermoTop prevents condensate formation for more accurate results.

www.eppendorf.com/thermomixer-c

Eppendorf®, the Eppendorf logo and Eppendorf ThermoTop® are registered Trademarks of Eppendorf AG, Germany.
Eppendorf ThermoMixer™, Eppendorf SmartBlocks™, Eppendorf ThermoStat™ and Eppendorf QuickRelease™ are Trademarks of Eppendorf AG.
All rights reserved, including images and graphics. Copyright © 2013 by Eppendorf AG.



No holds barred. *Science* Editor-in-Chief Bruce Alberts moderated the debate.

SCIENCE DEBATE

Scientists Clash Swords Over Future Of GM Food Crops in India

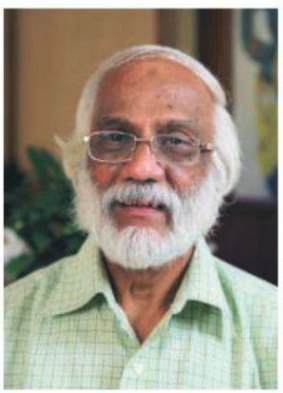
HYDERABAD, INDIA—One of the most contentious issues roiling India these days is whether the country should permit commercial planting of genetically modified (GM) food crops. A defining moment in the debate came in February 2010, when Jairam Ramesh, then–minister of environment and forests, called for a moratorium on the cultivation of brinjal, or eggplant, engineered with a gene from the bacterium *Bacillus thuringiensis* (Bt) that codes for an insect-killing toxin (*Science*, 12 February 2010, p. 767). The previous year, India's top biotechnology regulatory body had concluded that Bt brinjal is safe for environmental release. Public hearings held across India to discuss that recommendation tapped deep unease over GM foods. In response, Ramesh announced the ban on Bt brinjal, which he said would remain in effect until studies establish "the safety of the product from the point of view of its long-term impact on human health and [the] environment."

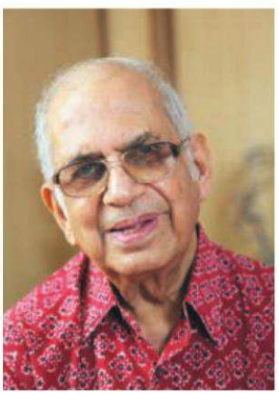
Watch the dabate at **www.scim.ag/vidGMO**

Three years later, the moratorium's repercussions are still being felt. Although the ban did not target research, Indian biotechnologists say that they have had a difficult time getting funding for GM experiments and permission for field trials (*Science*, 17 August 2012, p. 789). Critics and backers of the technology agree on one point: India's rules for regulating GM crops must be strengthened. A bill introduced in Parliament last month aims to do just that by setting up an independent Biotechnology Regulatory Authority of India to assess the safety of genetic modification. All eyes are now on the Indian Supreme Court, which is mulling a petition filed by activist groups demanding a prohibition on GM food crops in India; on 17 October 2012, a panel appointed by the court to advise it on the case recommended a 10-year moratorium on the introduction of GM crops. A decision is expected in the coming weeks.

Science sought to shed light on the issues by bringing together two prominent voices in the scientific community to debate the future of GM food crops in India. Speaking for the technology's backers was G. Padmanaban, a biochemist and former director of the Indian Institute of Science in Bangalore. Speaking for opponents of GM food crops was Pushpa M. Bhargava, a biochemist and former director of the Centre for Cellular and Molecular Biology (CCMB) in Hyderabad. The debate was moderated by Bruce Alberts, *Science*'s editor-in-chief, and held here at CCMB on 4 April. What follows is an edited excerpt.

–PALLAVA BAGLA AND RICHARD STONE





Pro and con. G. Padmanaban (*left*) and Pushpa M. Bhargava found little common ground on whether to commercialize GM food crops.

CREDITS (TOP AND BOTTOM): PALLAVA BAGLA (3)

B.A.: What is your impression of the indefinite moratorium imposed on the release of Bt brinjal for commercial cultivation?

P.B.: I believe that the indefinite moratorium that was put on open release of Bt brinjal was perfectly justified, because people did not want it. Jairam Ramesh had about seven or eight public meetings spread all over the country, and the overwhelming opinion was that it will not be in the interests of people in India to have the cheapest vegetable which is available all round the year, that is brinjal, to be genetically engineered, and that genetically engineered brinjal be available without labeling, for consumption by people. And they felt they had the right to decide what they were going to eat and what they will not eat.

G.P.: I believe this moratorium was very unfortunate. Actually, Bt brinjal was thought in terms of demonstrating a proof of principle so far as a food crop is concerned. I personally believe India would need Bt rice at some point of time. So this moratorium has sent a very wrong signal, in my opinion. That decision was more populistic than based on science as such. And it has depressed most of the scientists in the area. This is something which the country should worry about. People in this field have lost enthusiasm. Even students are not willing to get into this, which I think is very, very unfortunate.

B.A.: In an article in the December 2012 issue of *Frontiers in Genetics*, M. S. Swaminathan, distinguished leader of the green revolution in India, begins with the following statement: "I believe that the current concerns of biosafety and the impact of GMOs [genetically modified organisms] on biodiversity will soon give way to an appreciation of the potential benefits that the new genetics can confer on humankind." Do you agree or disagree with that statement?

G.P.: I personally believe this biodiversity card is overplayed. After all, you will see genes have been transferred vertically, hori-

zontally, all through evolution. For example, if you look at the rice genome, how many fungal genes are there, how many viral genes are there, how many bacterial genes are there? There is nothing like a pure rice genome. So to think a couple of genes would alter the biodiversity, I really do not buy that argument because in nature every plant has been modified. The only concern in my perception is whether the gene we are introducing is safe enough.

Of course, safety is a prime concern. I have no argument on that and safety is needed. Take Bt as an example. Millions of people have been consuming Bt corn for over 15 years—Americans, Canadians, Chinese, South Africans, Argentineans, Brazilians—and I have not seen any authenticated report of any environmental risk or health risk so far as this is concerned. Bt brinjal was 8 years in trials. It was not as if overnight somebody decided that Bt brinjal should come in. Many scientists were involved in this process.

B.A.: Swaminathan was saying here that the current concerns will disappear soon, and you certainly agree with that.

G.P.: Yeah, I definitely agree with that. Current concerns, I hope, will disappear but there is I feel a deliberate attempt in India to keep raising these concerns.

B.A.: Dr. Bhargava?

P.B.: Well, as far as Swaminathan's statement is concerned, I think it is a very neutral statement that when these concerns will cease to exist, that may take 50 years, that may take 100 years, that may take 5 years.

As regards the other issues that my friend Padmanaban has raised: There is a great deal of evidence that there have been health problems amongst Americans, especially related to allergy, since the large-scale consumption of Bt corn or GM corn and GM soya started in the U.S. In fact, if you plot qualitatively the increase in incidence of gastrointestinal disorders amongst Americans over the last 12 to 13 years and the increase in the consumption of GM food, the two curves seem to overlap very substantially. And there is evidence in Latin America and Brazil where there has been increased consumption of GM crops, that there is an increase in incidence of childhood cancer and several other problems. So to say that there is no evidence of any deleterious effect on human health, on animal health, on plant health, and on biodiversity … I think is ignoring a tremendous amount of evidence that these effects are very real.

CREDIT: DENNIS COOK/AP PHOTO

CLIMATE CHANGE

Hansen's Retirement From NASA Spurs Look at His Legacy

For decades, American climate scientist James Hansen published important papers on global warming and shared his data at influential congressional hearings—along with his policy prescriptions. He tussled with White House officials over his right to speak his mind, lobbied leaders the world over, and testified in defense of jailed activists. The 72-year-old has also been arrested five times in protests against the continued burning of fossil fuels or to demand that the United States put a price on carbon emissions.

Few other figures in modern science have straddled—and for that matter blurred—the boundaries between science, policy, and advocacy quite like the homespun but outspoken climatologist. Now, with his 2 April retirement announcement from NASA's Goddard Institute for Space Studies (GISS) in New York City, where he served as director, Hansen is entering a new and perhaps final phase of a storied career. He wants to continue publishing as an independent scientist (although funding is proving tough) and ramp up his activism. The move has helped highlight a long-simmering debate: Is Hansen a role model to be emulated by younger researchers—or a polarizing figure whose tactics have proved counterproductive?

"He has done very important science really well," says Michael MacCracken of the Climate Institute in Washington, D.C. "[And] for those whose scientific findings relate to environmental and societal welfare, Jim has been demonstrating the additional obligations that come with doing scientific research in the public service."

Hansen is "among the best climate scientists," agrees Ken Caldeira of the Carnegie Institution for Science in Palo Alto, California. But "it's important to keep value and opinions separate from scientific judgments about empirical fact and, especially in the last 5 years, [Hansen has] not made clear enough distinctions."

In an e-mail to some 7000 recipients of his regular missives, Hansen explained last month that "my aim in 'retiring' is to have more time to focus on science, to try to make the science clearer to the public, and to connect the dots all the way to policy implications." And in a 4 April editorial in the *Los Angeles Times* opposing the construction of the Keystone XL oil pipeline from Canada to the United States, Hansen did just that. "The perspective of pipeline apologists is contrary to the laws of physics and basic economics, neither of which gives a damn about politics," he wrote.

It's the kind of rhetoric that has made Hansen a media favorite. As a scientist, however, he began his career far from the hot lights

Testify. James Hansen's iconic congressional testimony in 1988 brought the global warming issue to the national stage.

of the TV studio. He joined NASA in 1967 as an astrophysicist studying Venus, later turning his focus to the greenhouse effect on Earth. National acclaim came after he appeared at blockbuster hearings held by the U.S. Congress in 1988, where he declared that the world was warming, that humans were most likely responsible, and that his climate models suggested that heat waves would increase as a result. "It is time to stop waffling," he argued, and his testimony appeared on the front page of newspapers around the world.

Many scientists, however, were put off. "What really bothers them is not that they believe Hansen is demonstrably wrong, but that he fails to hedge his conclusions with the appropriate qualifiers that reflect the imprecise science of climate modeling," *Science*'s Richard A. Kerr reported a year later (*Science*, 2 June 1989, p. 1041).

Over time, however, Hansen's biggest pronouncements have proven essentially correct as the data have come in. "When I saw him give one of his plenary talks at the [American Geophysical Union], I would guess his estimate of [the planet's sensitivity to carbon] is overconfident, he did a very clever job making the case for sensitivity near 3° while staying in line with the facts as I understand them," recalls climate expert David Keith of Harvard University. Still, some of Hansen's views place him at a far end of the spectrum of the scientific consensus on climate change; he's had to qualify, for instance, some of his previous statements about the risk that Earth could spiral into a steamy, Venus-like environment. And a paper published last year in the *Proceedings of the National Academy of Sciences*, arguing that recent heat waves in Russia and the U.S. West were "a consequence of global warming," drew fire from scientists who published a competing analysis that found no direct connection.

Hansen's "exceptional trust in physical intuition," as Kerr called it, along with his determination to be heard, complicated his role as a government scientist. As GISS director in 1989, he scuffled with the administration of President George H. W. Bush over caveats about climate models that bureaucrats in the White House's Office of Management and Budget (OMB) wanted to add to his congressional testimony. Hansen argued the additions would suggest his warnings were overblown. "I should be allowed to say what is my scientific position," he added at the time. "There is no rationale by which OMB should be censuring scientific opinion." Hansen used much the same language when, 16 years later, he fought with the administration of President George W. Bush, which attempted to bar him from giving certain interviews; Hansen won after he leaked the skirmish to the press. Such public feuds, and his outspoken views, may have led the Department of Energy to pull funding from GISS in the 1980s, Hansen told *Science* last summer. And the "sideshow" that his high profile sometimes created was growing difficult lately for NASA leaders to countenance, he added. Still, many of his GISS co-workers have said they believe that Hansen's work only helped burnish the small institute's reputation.

In protest. Hansen's willingness to be arrested as part of demonstrations against the burning of fossil fuels sets him apart.

Hansen's impact on policy also draws divided reviews. At first, his voice had big resonance, says Rafe Pomerance, who worked on climate policy as an environmental activist, Congressional staffer, and executive branch official. "After his '86 testimony, the [U.N. Intergovernmental Panel on Climate Change (IPCC)] was set up. After his testimony in '88 and '91, the U.N. Framework Convention on Climate Change got started," Pomerance notes. But Hansen's later forays into policy during the George W. Bush and Obama administrations may have had less relevance—in part because by then Hansen had become just one of many voices on the issue, "diluting" the potency of his views, says Roger Pielke Jr., a policy expert at the University of Colorado, Boulder.

Still, Hansen's prominence allowed him, as a private citizen, to arrange lobbying visits with world leaders that few rank-and-file government scientists could dream of. Such independent advocacy "would be unheard of if he was, say, a diplomat or CIA agent freelancing on some issue outside of government policy," Pielke says. Others, though, believe that it was an appropriate role. "It's disheartening that he has to [now] remove himself from a federal position to advocate on climate change. Government exists, in theory at least, to serve the public's best interests," says Emmy Burns, a student activist at the University of Wisconsin, Madison. Now, "the question is will he have as powerful a role as an advocate?" Pomerance says. "We'll see."

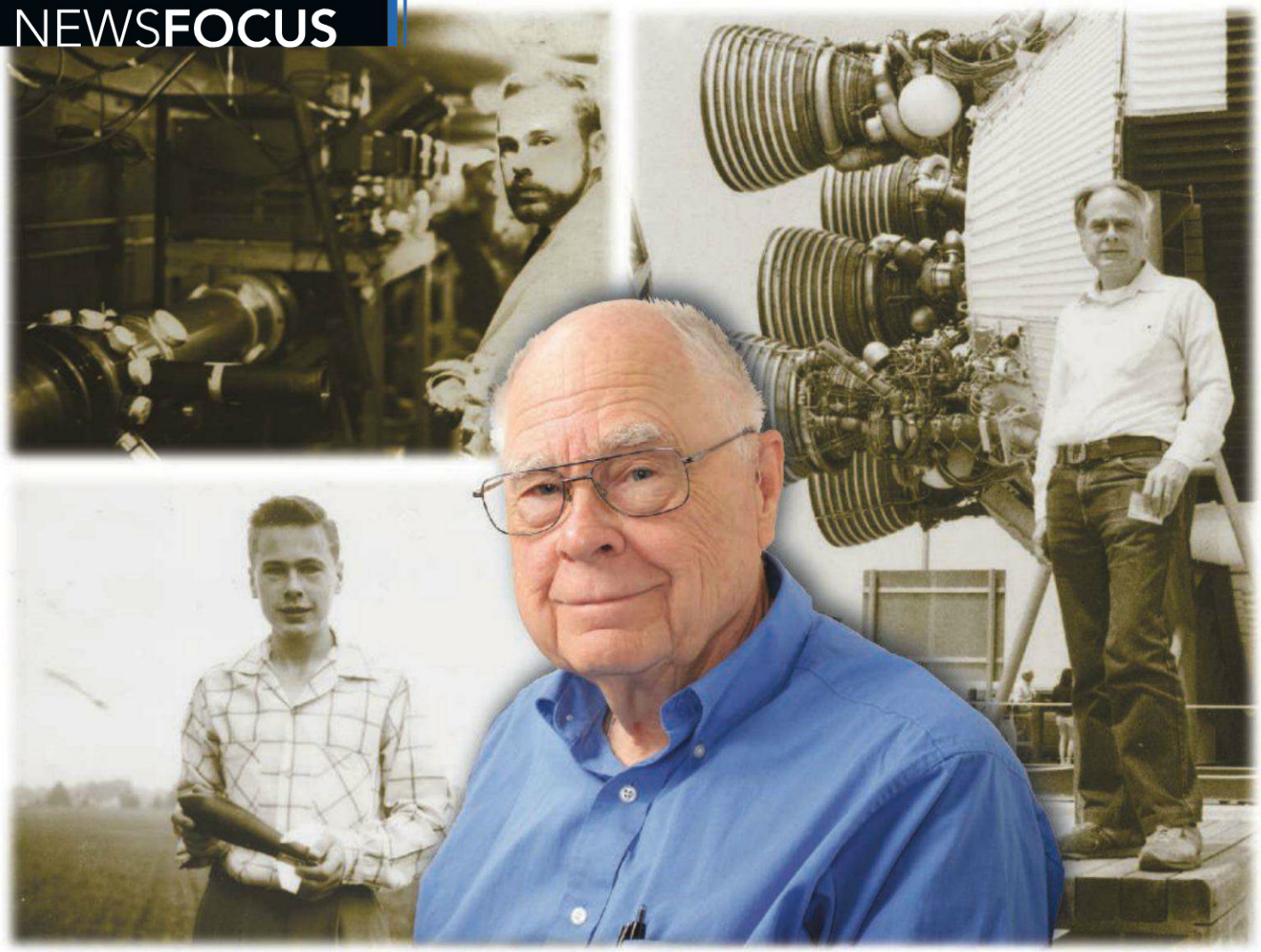
Hansen's influence with other researchers is more difficult to assess. In the late 1980s, "there were a lot of scientists who were not paying attention to this issue," Pomerance says, and the 1988 hearings helped change that. But few other researchers have followed Hansen's example and been willing to risk arrest at public demonstrations, or play such a vocal role in public debates. Hansen has also never formally contributed to the IPCC process, researchers note, underscoring his loner status among his peers. But Juliette Rooney-Varga, a marine biologist at the University of Massachusetts, Lowell, says "as a scientist and a mother of three young children, I am grateful for the example that he has provided. … Hansen took substantial professional risk in choosing to act on his moral conscience, while still maintaining an impressive scientific career."

Hansen says that he'll keep doing research despite his departure from NASA. He's trying to create a small new institute with key GISS scientists, but fundraising for that effort has proven "difficult and time consuming," he tells *Science*. In the meantime, he's keeping up his outreach efforts, including a book that he's writing called *Sophie's Planet*, consisting of letters between him and his granddaughter.

Many observers are convinced that Hansen will remain a force to be reckoned with. "He's sui generis—that's for sure," says marine biologist James McCarthy of Harvard. "He's been ahead on the science for decades and has played a very important role in communicating the science of climate change to the public. … I've disagreed with him on some of his views on policy, … but let's be clear: The world needs Jim Hansen."

–ELI KINTISCH

CREDIT: BEN POWLESS/WIKIMEDIA COMMONS



Mr. Borucki's Lonely Road to the Light

The father of NASA's Kepler orbiting exoplanet finder had to pioneer new optical techniques and overcome decades of skepticism to get his pet project off the ground

At 74, many men are happy to while away their time gardening or playing poker or watching reruns of old sitcoms. William Borucki, the architect and principal investigator of NASA's exoplanet search mission, Kepler, is cut from a different cloth. On weekends, he likes to take off with his wife, Josephine, and his boyhood buddy, Gene Westerberg, to trek through Alum Rock Park, a few miles east of San Jose, in search of kempite, a rare manganese mineral found nowhere else on Earth. So far, Borucki hasn't found any, and perhaps he never will. It doesn't matter. Borucki's journey is his destination.

If he didn't enjoy the journey so much, Borucki would have given up long before he realized his goal of launching a spacecraft to find planets outside the solar system. Kepler, which has opened the floodgates for exoplanet discoveries since its launch in 2009, would never have come about. Those who have followed Kepler from idea to reality say that the mission is a testament to Borucki's ingenuity and vision and iron will, a spirit so inured to rejection and failure that it might as well be wrapped in rhino hide.

In the late 1980s, Borucki proposed the idea of finding exoplanets by measuring the dip in a star's brightness when an orbiting planet travels across the face of the star. For several years, he was alone in pushing the idea against a tide of scorn and hostility from many in the field who did not believe that the technique could work.

"There was no one from NASA headquarters who would support him," says David Morrison, an astronomer at NASA Ames Research Center near Mountain View, California, who was Borucki's supervisor during the early years of the struggle. "It took a lot of courage on his part because of all the negative reactions." In the end, Morrison says, Borucki proved to be "the fighter who is knocked down and gets up, is knocked down and gets up again." For this perseverance alone, some might say, Borucki should be awarded a Ph.D. That way, he'll no longer have to correct those who refer to him as Dr. Borucki.

An urge to explore

Borucki is 5'6" with kindly eyes and bushy eyebrows. One afternoon last November, he sat down for lunch with Westerberg and Josephine in the outdoor area of a restaurant, under the warm California sun. As he dug

CREDITS: (BACKGROUND) COURTESY OF WILLIAM J. BORUCKI; (FOREGROUND) DOMINIC HART/NASA AMES RESEARCH CENTER

Rocket man. (*Clockwise from lower left*) Borucki as a teenaged model rocket buff in Wisconsin; at an underground ballistics range in 1962; at Kennedy Space Center with a retired Saturn V engine assembly in 1985; and today (*large image*).

into his salad, Westerberg recalled growing up with Borucki in Delavan, Wisconsin, where Borucki's father worked at a factory that made clocks for automobiles.

Tinkerers since childhood, Borucki and his brother built rockets when they were in high school, cutting the tips off matchsticks for fuel. Later, they and Westerberg switched to homemade gunpowder, crushing an artists' drawing set for charcoal and furtively buying the other ingredients from the town's two drugstores. "One of us could go and get the sulfur from one store, and the other could get the saltpeter from the second store," Westerberg says. "We didn't want people to connect the two."

Building rockets was just one of many pursuits. "We built a jet engine in the basement of our house," Borucki recalls. "It worked, but it burned off part of the paint off the wall. It meant you had to clean things up pretty quickly before your folks came home." Borucki also got interested in astronomy during those years, biking to the Yerkes Observatory 15 miles away to look at the heavens. He would climb up on the garage of his house and lie on the roof to watch meteor showers. Looking up at the sky, Borucki says, he knew what he wanted to do most of all was "to go out and explore the galaxy."

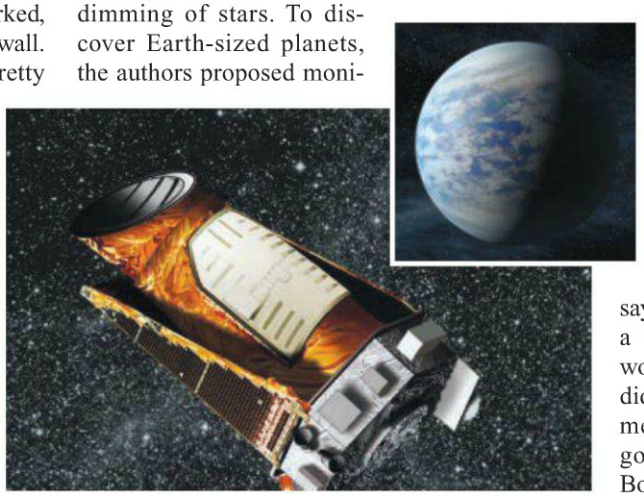
Borucki studied physics at the University of Wisconsin, Madison, earning a master's degree in 1962. Pursuing a Ph.D. didn't seem appealing. "I really don't like to just think about things," he says. "I like working on things. I like to do things. I like to see an idea put into practice."

Borucki went to work at NASA Ames and spent a decade developing heat shields for the Apollo program. In 1972, when Apollo ended, NASA fired everybody working on the program ("No good deed goes unpunished," Borucki jokes) but let them apply for other positions at the agency. He found another spot at Ames as a researcher in space sciences.

For the next few years, Borucki studied the Earth's atmosphere and lightning on other planets in the solar system. He also attended seminars at which scientists presented ideas about ways to find planets outside the solar system. The discussions rekindled Borucki's boyhood dreams of galactic exploration.

The seminars focused on a technique called astrometry: detecting a planet by measuring the slight wobble the planet's gravitational tug gives to its parent star. Borucki was drawn to another approach: measuring the dip in the light of a star when a planet passes in front of it. A colleague referred him to a paper that an artificial-intelligence researcher named Frank Rosenblatt had published in *Icarus* in 1971. The paper was the first to propose transit-based detections of planets.

Borucki tried to contact Rosenblatt, who had spent years running computer simulations at the Cornell Aeronautical Laboratory, but learned that Rosenblatt had died in a canoeing accident shortly before the *Icarus* paper appeared. Borucki picked up the concept from where Rosenblatt had left it. In 1984, he and his Ames colleague Audrey Summers published a paper that laid out the potential for discovering Jupiter-sized planets with ground-based telescopes using high-precision light detectors to track the periodic dimming of stars. To discover Earth-sized planets, the authors proposed monitoring thousands of stars simultaneously from above Earth's atmosphere.



Like home. Kepler with Kepler-63c (*artist's conception, inset*), a "super-Earth" detected in the habitable zone of a sunlike star.

The paper got a lukewarm reception. "People generally ignored it," Borucki says. The next year, he published another paper with a colleague predicting that the natural variability of stars would limit researchers' ability to pick out the change in brightness caused by a transit, unless more advanced photometers were developed. "That paper was ignored, too," he says.

But Borucki pressed ahead. Using small grants of $10,000 to $15,000 from the Director's Discretionary Fund at Ames, he organized workshops to hash out how to build precise photometers. With colleagues at Ames and elsewhere, he began building prototypes that could come close to measuring the dimming of starlight with the necessary precision: a 1% dip in brightness to detect a Jupiter-sized planet around a sunlike star, a 0.01% dip to detect an Earth-sized planet.

"Here's Bill again"

Even as Borucki made steady progress in building high-precision photometers, he struggled to win support for his larger concept: monitoring thousands of stars across a wide field of sky to look for periodic dimming. "Here was this young guy with a crazy idea," Morrison says. "He had no history of building telescope detectors. He was just not part of the astronomy club. I think people probably didn't trust him." Morrison says that Borucki's presentations at scientific meetings drew mildly hostile reactions. "It was, 'Oh, here's Bill again. We got to listen to him for half an hour, and then we can go back to more useful things like astrometry.' "

Not having a Ph.D. didn't help, Borucki says. He recalls wincing when a Harvard University professor at Ames presented results on atmospheric chemistry: The professor shrugged off an unexpectedly low reading of stratospheric water vapor by remarking that the researcher who made the measurements didn't have a doctorate and had probably messed up. " 'Thanks a lot,' " Borucki recalls saying to himself. " 'You don't know a damn thing about experimental work, and just because this person didn't have a Ph.D., you doubted his measurements?' I have never forgotten that." On another occasion, Borucki says, officials at a university where he was slated to teach for a year withdrew his invitation on learning that he wasn't Dr. Borucki.

Morrison himself was skeptical about transit searches when he became Borucki's supervisor in 1988. A year later, he arranged for a panel of space scientists to review Borucki's idea of finding transiting planets by simultaneously observing thousands of stars with a charge-coupled device (CCD) detector. After Borucki's presentation, the panel, led by astronomer Jill Tarter, grilled him for several hours in the equivalent of a marathon dissertation defense.

"I had an awful lot of things to defend," Borucki says. Although CCDs were becoming increasingly popular, nobody had yet shown that they could work at the level of precision that Borucki's concept demanded. The panel also wondered whether the intrin-

CREDITS: NASA/AMES/JPL-CALTECH (2)

eppendorf & Science

PRIZE FOR NEURO BIOLOGY

2012 Winner
Dr. Marlene R. Cohen
Assistant Professor
University of Pittsburgh

Application Deadline
June 15, 2013

Call for Entries

Eppendorf & Science Prize for Neurobiology
The annual Eppendorf & Science Prize for Neurobiology, an international award, honors young scientists for their outstanding contributions to neurobiological research. The winner and finalists are selected by a committee of independent scientists, chaired by Science's Senior Editor, Dr. Peter Stern. To be eligible, you must be 35 years of age or younger.

You could be next to win this prize and to receive
> Prize money of US$25,000
> Publication of your work in Science
> Full support to attend the Prize Ceremony held in conjunction with the Annual Meeting of the Society for Neuroscience in the USA
> An invitation to visit Eppendorf in Hamburg, Germany

It's easy to apply!

Learn more at: www.eppendorf.com/prize

eppendorf

Science AAAS

eppendorf® and the eppendorf Logo are registered trademarks of Eppendorf AG, Hamburg, Germany. All rights reserved, including graphics and images. Copyright © 2013 by Eppendorf AG.
AAAS® and Science® are registered trademarks of the American Society for the Advancement of Science, USA. Kurstin Roe Photography.

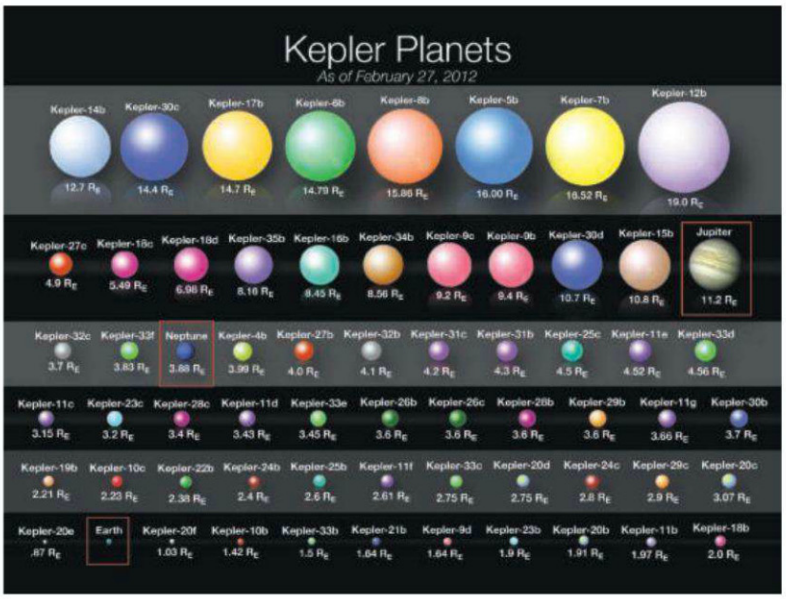
sic variability of stars—still unknown at the time—would prove too high for Borucki's proposed detectors to pick out the signature of a transiting planet. Borucki argued convincingly that the variability in stellar brightness ought to resemble that of the sun—a manageable amount of "noise" for his proposed instruments to handle. At the end of the day, the panel gave him a thumbs-up. When NASA announced a new program called Discovery to support intermediate class space missions, Borucki seized the opportunity. He persuaded a handful of colleagues at Ames, including a physicist named David Koch who was exceptionally talented at building instruments, to join him in proposing a mission concept clunkily named the Frequency of Earth-Sized Inner Planets. It was 1992—the year that astronomers detected the first exoplanets, circling a stellar cinder called a pulsar.

The proposal was rejected. Reviewers weren't convinced that light detectors then available could achieve the required precision. In 1994, Borucki and his team tried again, bolstering the proposal with new studies to show that CCDs could do the job. This time, reviewers rejected the idea as too expensive. In 1995, astronomers discovered the first exoplanet orbiting a sun-like "main sequence" star. "We felt it would help us because now we could show that there were more planets to be found," Borucki says. But the team's next submission—now named Kepler—was rejected again in 1996. The reviewers pointed out that nobody had yet monitored the brightness of thousands of stars simultaneously. "They said, 'Build an observatory and show that it can be done,'" Borucki says.

To space at last

One day in 1996, Borucki drove half an hour from Ames to the University of California's Lick Observatory to inspect one of its small domes. The telescope in it hadn't been used in years. The floors were rotting, the roof leaked. But Borucki was smitten. "It was beautiful," he says. "It said *astronomy*. It said *here's a machine that you can use to build your instrument*. In reality, of course, the dome didn't rotate; there was no place for a computer, and you'd freeze to death. But don't you see, I helped my dad build a house. What's so hard about building a floor?"

Over the next several months, Borucki and his colleagues worked to overhaul the dome, often spending nights there in the company of rattlesnakes and rats. Everything had to be done on a shoestring budget. After the team had built the photometer and hooked it up to the telescope, Borucki found a way to calibrate the detectors on the cheap, using a Teflon bucket that he had bought at a hardware store. He put a bulb in the bucket and put the lid back on, using it as a diffuser to uniformly spread out the light. Hoisting the bucket on a metal pole, Borucki tipped it horizontally and pointed the lid at the telescope. It was a simple idea that worked perfectly, says Natalie Batalha, a member of the Kepler team. By the end of 1997, Borucki and his colleagues were successfully monitoring the brightness of some 6000 stars simultaneously.



Sampler. Kepler's discoveries include exoplanets about the size of Earth.

Armed with the new data, the team made a fresh bid in 1998 to get the mission greenlighted but met with another rejection. Reviewers were concerned that the detectors wouldn't be able to pick out signals of transiting Earth-sized planets amid the noise expected in orbit from vibrations experienced on a spacecraft. Borucki reacted to the rejection—the fourth one in 6 years—with his typical sequence of frustration followed by new resolve. "When I get angry, I usually rant for a while," he says. "You don't want to be around me for a couple of days. I rant for a while and then I say, 'Well, rant's done. Time to get the job done.'"

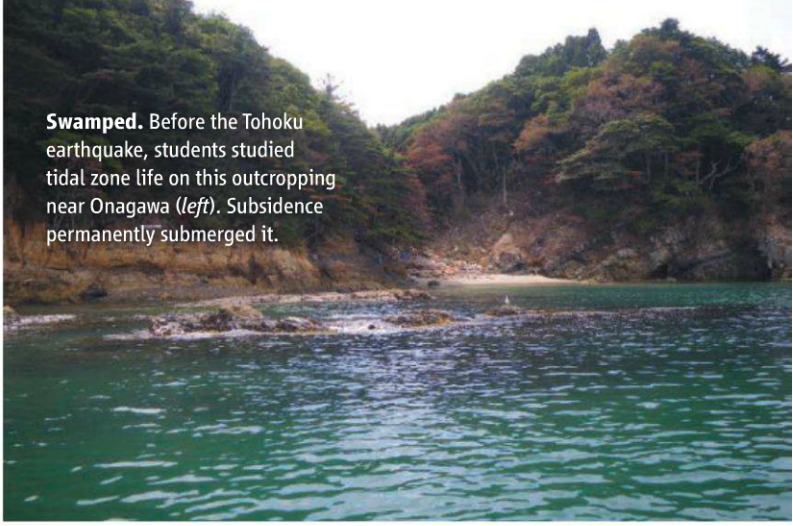
With $500,000 from NASA Headquarters and $500,000 from NASA Ames, Borucki and his colleagues—including Koch and another physicist named Fred Witteborn—built a prototype of the instrument that they were proposing to launch into space. It was a metal box the size of a refrigerator with a small telescope inside pointed at a ball with tiny holes drilled into it. When the inside of the ball was illuminated, the holes resembled pointlike stars as would be seen by the telescope. A fine wire was strung across each hole; when a current was passed through it, it expanded slightly, blocking more light than normal to simulate the dimming of a star caused by a transit. "They [NASA reviewers] were assuming that it would take us a few years to complete this step, but we worked as fast as we could," Borucki says. In 2000, the mission was finally approved. It cost $640 million and 9 years of design, building and testing to finally launch the mission in 2009.

In the 4 years that Kepler has been in space, monitoring the brightness of some 150,000 stars, it has discovered more than 2700 possible exoplanets. At last count, follow-up observations by ground-based telescopes had confirmed 122 as planets. Roughly half of the candidates are estimated to be twice the size of Earth or smaller; many of these could be rocky planets. The findings suggest that hordes of Earth-sized, Earth-like planets may lurk in the habitable zones of stars, waiting to be discovered.

Borucki is convinced that astronomers will find mirror Earths in the not too distant future. And although budget constraints have made many researchers pessimistic about the prospects of flying a next-generation observatory to study such planets, Borucki is confident that such a mission will fly sooner or later. The first goal will be to study the composition of these planetary atmospheres in detail. "I don't know when such a mission will fly or if it will be a U.S. mission or a Chinese mission or a European mission," he says. "But that mission will fly. That's a certainty. You can predict some of these things. There are things that have to happen. They must happen." Like Kepler.

–YUDHIJIT BHATTACHARJEE

CREDIT: NASA/KEPLER MISSION/WENDY STENZEL



Swamped. Before the Tohoku earthquake, students studied tidal zone life on this outcropping near Onagawa (*left*). Subsidence permanently submerged it.

Taking the Pulse of a Ravaged Ocean

The 11 March 2011 earthquake and tsunami wreaked havoc on fisheries and the marine environment. Researchers want to document—and aid—the recovery

SHIZUGAWA BAY, JAPAN—The engine cuts to an idle as the fishing boat reaches its target coordinates. After an "OK!" from the crew chief, the deck explodes with activity. Off the starboard rail, crew members drop 10-liter plastic bottles attached to ropes and reel them in by hand to collect samples from the surface, then from 10, 20, and 30 meters down. At the port side, the crew chief casts a 3-meter-long fine mesh plankton net. After a few minutes, he gathers it in by hand and funnels the catch into small jars. To measure turbidity, he sinks a tethered white acrylic disk, the size of a dinner plate, and notes the depth at which it disappears from sight.

Prior to 11 March 2011, this bay on Japan's Pacific coast, east of Sendai, was famous for wakame seaweed and oysters cultured on ropes and racks suspended from buoys. The massive tsunami unleashed by a magnitude-9.0 earthquake that day swept all that away, as well as most of the fishing boats and onshore facilities.

Buoys once again dot the bay, indicating a return of aquaculture. But it is by no means back to business as usual. Last year, the normally deep-green wakame was yellowish, and nobody wanted to buy it. Solving the mystery of the pallid seaweed is one of several challenges facing the scientists. Abalone and sea urchin, once plentiful in shallow waters, have disappeared. And fin fish are scarce. Fishing is bringing in just 20% to 30% of predisaster revenues in a region whose economy depends on the bounty of the sea.

This cruise on a crisp March morning is part of an unprecedented effort to understand how the Tohoku tsunami affected marine ecology and to monitor its recovery. Some 300 researchers from more than a dozen institutions will spend the next decade documenting everything from changes in water quality, pollution, and the strength of currents to shifts in biodiversity and the population genetics of marine life. One big challenge is teasing out how intensive coastal development and wetland reclamation may have exacerbated the marine disaster, and how onshore rebuilding may degrade the marine environment. Working with local fishing cooperatives, researchers intend to help develop sustainable fisheries plans and revitalize the region's aquaculture.

Beneath the surface

Smashed buildings, wrecked cars, and washed-out roads and railways, not to mention a horrific loss of life: The devastation visited on coastal communities by the earthquake and tsunami, which topped 3-story buildings in hard-hit locations, was obvious. The disaster's effects on marine ecosystems are also severe, but they are harder to discern. Tectonic shifts reshaped the coastline, submerging wetlands and altering currents. Turbulence destroyed kelp forests and washed away shellfish beds and fish breeding grounds.

Tsunamis are a regular phenomenon, and coastal environments recover over time. But the March 2011 disaster was unique. According to Akihiro Kijima, a marine population geneticist at Tohoku University in Sendai, the amount of humanmade debris flushed out to sea by the Tohoku tsunami vastly exceeded that of the December 2004 Indian Ocean tsunami. The Tohoku coast was far more developed than coastlines hit in 2004, with the narrow shore thick with buildings and industrial facilities. Besides sweeping out bits and pieces of houses, smashed boats, and cars, fuel oil and chemicals leaked from storage tanks and factories. And, in a first, the tsunami resulted in massive radiological contamination of the ocean (see sidebar, p. 547).

Launched in January 2012, the 10-year Tohoku Ecosystem-Associated Marine Sciences (TEAMS) project is still establishing observing sites and gathering baseline data. In some locations, researchers have been monitoring water quality and marine life for years; in others they will have to estimate and extrapolate to predisaster conditions.

But scientists are already making preliminary observations. One is the likely reason for the yellowed wakame. Kijima, one of the project leaders, blames coastal subsidence, which appears to have altered currents. Fresh ocean water no longer sweeps into the inner reaches of Shizugawa Bay, and turbid water is not being flushed out, so near-shore marine life forms are starved of nutrients, he says.

Netting results. Plankton abundance is a key benchmark of ocean health.

CREDITS (CLOCKWISE FROM TOP LEFT): M. IKEDA; M. KANNO; D. NORMILE/SCIENCE

The Pacific Swallows Fukushima's Fallout

The Tohoku disaster had one unprecedented impact: an enormous amount of radiation was deposited in the ocean. The good news is that any long-term effects are likely to be trivial. And, thanks to strict testing of seafood, the radiation should not harm human health. But there are lingering conundrums.

Flooding from the tsunami overwhelmed the Fukushima Daiichi Nuclear Power Plant, triggering hydrogen explosions on 12 and 14 March 2011; prevailing winds blew most of the vented radiation out to sea. A second calamity occurred from 2 to 6 April that year, when an estimated 520 cubic meters of highly contaminated cooling water leaked into the ocean. Together, the two events deposited an estimated 3.5 to 27 petabecquerels into the Pacific—a small fraction of the radiation released during the Chernobyl accident.

Measured radioactivity levels in seawater fell rapidly, thanks to dilution and radioactive decay. At many monitoring points, radiation levels have returned to background levels, says Masashi Kusakabe, a chemical oceanographer at the Marine Ecology Research Institute in Tokyo, which has monitored the ocean near Fukushima for almost 30 years.

One puzzle is a small harbor formed by a breakwater in front of the Fukushima plant. Once the cooling water leak was plugged, radiation levels there declined. But they remain elevated above those of the nearby ocean, suggesting that "there is some radioactivity coming from somewhere," says Jota Kanda, a biogeochemical oceanographer at Tokyo University of Marine Science and Technology. Tokyo Electric Power Co., the plant operator, has maintained that there is no evidence of continuing leaks from the power plant to the sea.

Hot water. Elevated radioactivity persists in the sea and sediments near the Fukushima Daiichi Nuclear Power Plant, but in puzzling patterns.

There are other anomalies among the 30 seawater locations monitored by the Marine Ecology Research Institute. Two radiation hot spots are close to the mouth of the Naka River, 125 kilometers south of the power plant; another one is 70 kilometers north of the plant, near the outlet of the Abukuma River. Kusakabe says it is possible the two rivers are carrying fallout-laden sediment into the ocean. But there is another location with a persistently high level of radioactivity 75 kilometers south of the plant that is not near any major river, and points closer to the power plant are less contaminated. "We don't have an explanation for this," he says.

Contamination in marine sediment near Fukushima is even more troubling. The seafloor is a mosaic of muddy, sandy, and rocky patches, and currents move sediment to and fro, Kusakabe says. From the surface, he says, it is difficult to ensure that grab samplers hit the same spot for comparisons over time. Despite these challenges, the team has measured levels as high as 300 becquerels per kilogram of radioactivity in sediment, far higher than the 1 becquerel per kilogram measured prior to the accident. And some plankton and fish samples have also turned up with unexpectedly high levels of contamination.

None of this threatens human health, Kusakabe says. Radioactivity levels in seawater are below the limit set in Japan for drinking water: 100 becquerels per liter. And marine products from Tohoku are tested before going to market, with anything over 100 becquerels per kilogram rejected. Fishing is still prohibited off Fukushima, and it's unlikely to resume until the riddle of the contaminated sediment is solved. **–D. N.**

At least one group was off the blocks before TEAMS started. Just 3 months after the disaster, Tohoku University's Masakazu Aoki, who studies benthic organisms, began monthly monitoring of 26 sites in three bays. His group noted that turbulence from the tsunami washed away algae, such as kelp, or stripped off its fronds. And many areas of the rocky bottom were smothered by mud. Abalone, sea urchins, and other organisms that feed on the algae largely disappeared, except for some mature individuals that somehow hung on.

By last fall, things were looking up. Aoki's team observed new kelp growth, although kelp in water deeper than 4 meters was faring poorly because of persistent turbidity. Juvenile sea urchins were also reappearing. Curiously, abalones were still missing, and it is not clear why, Aoki says.

He and his group found that the harm from the earthquake and tsunami extends to areas that escaped the brunt of the waves. The west side of Oshika Peninsula, which juts southward into the Pacific from a bend in Honshu Island, was sheltered from the tsunami, which struck from the east. But the coast in this area sank as much as 90 centimeters during the quake. Deeper patches of kelp were starved of sunlight and died. Although spared the turbulence of a direct hit, the west side of the peninsula was inundated with tsunami-dredged mud, which "had a profound effect on the rocky shore communities," Aoki says. There, as in waters battered by the tsunami, kelp, sea urchins, and abalones suffered greatly.

Changes to marine ecosystems may not be easily reversed. As part of TEAMS, Japanese researchers at a 23 March symposium in Sendai compared notes with scientists who studied the 2004 Indian Ocean tsunami. Zulfigar Yasin, a marine ecologist at Universiti Sains Malaysia, in Penang, reported that post-tsunami mud was a factor there, too. Total biomass recovered, Yasin says, but "there was a long-term change in the composition of fauna." Fish species that tolerate cloudy water moved in and are still there. Turbidity is expected to be a chronic problem in Tohoku, as dirt washes into rivers and bays during coastal reconstruction.

On Shizugawa Bay, mud is also on the minds of the sampling crew. Finished with the water column, they use a scooplike Smith-McIntyre bottom grab sampler to retrieve buckets of mud. From one grab, they carefully extract samples to preserve the order of sediment layers. From the next bucket, the crew chief notes the color and odor; he reassuringly reports that there is no oily or chemical scent. The team sieves the contents of several buckets for bottom-dwelling worms and crustaceans, and puts these in jugs. In 90 minutes, they have collected more than 70 samples of all sorts. Back in the lab, these will be analyzed for salinity, dissolved oxygen, chlorophyll, metallic and chemical contaminants, and the density and variety of organisms, among other things.

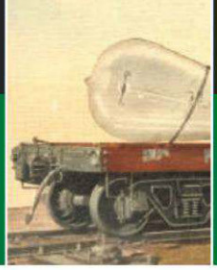
The captain revs the engine and points the boat toward a spot near the mouth of the bay. Findings from this cruise and others will not only aid in the region's recovery, Kijima says. Data will be archived, he says, "so it will be useful for the next marine disaster anywhere in the world."

–DENNIS NORMILE

CREDIT: MASNARI GENKO/*YOMIURI SHIMBUN*/ASSOCIATED PRESS

Manipulated images 551

Global research, inconsistent integrity 552



LETTERS | BOOKS | POLICY FORUM | EDUCATION FORUM | PERSPECTIVES

LETTERS

edited by Jennifer Sills

Geoengineering: Guidance Exists

IN THEIR POLICY FORUM "END the deadlock on governance of geoengineering research" (15 March, p. 1278), E. A. Parson and D. W. Keith recognize the environmental and policy risks posed by geoengineering methods and lay out a course of action to address them. They argue that a void in international governance of research exists and that geoengineering operations are subject to no international legal control. They further assess no progress in defining the boundary between small and large activities. In fact, relevant guidance does exist.

The 1977 Convention on the Prohibition of Military or Any Other Hostile Use of Environmental Modification Techniques—the so-called ENMOD Convention—is an international agreement that already governs aspects of geoengineering. The Convention was the result of a several-year policy development, set in motion by interest in the national security implications of geoengineering. The U.S. government investigated geoengineering techniques (*1*) and used weather modification (cloud seeding) as a weapon during the Vietnam War in the late 1960s and early 1970s (*2*). The Convention, recognizing the issue of scale, prohibits weapons uses that would have "widespread, long-lasting, or severe" effects. An agreed Understanding specifies that, for the purposes of the Convention, these terms refer to an area of several hundred square kilometers, a duration of months or about a season, and severity as "serious or significant disruption or harm to human life, natural and economic resources or other assets" (*3*). Parties commit to the "fullest possible" exchange of relevant information on peaceful geoengineering and to feasible international scientific and economic cooperation.

If geoengineering research, let alone experiments or possibly operations, is to proceed, the ENMOD Convention and its negotiating history provide precedents and information. The need for a moratorium on geoengineering for weapons purposes is moot. The authors' discussion of a social bargain recognizes the importance of "international norms of cooperation and transparency." Arms control already applies to geoengineering: Fortunately, one cooperative norm is already in place.

PIERCE S. CORDEN

Visiting Scholar, Center for Science, Technology and Security Policy, American Association for the Advancement of Science, Washington, DC 20005, USA. E-mail: pcorden@aaas.org

References

1. S. Weart, "The Discovery of Global Warming: Climate Modification Schemes" (2011); www.aip.org/history/climate/RainMake.htm.
2. D. Shapley, *Science* **184**, 1059 (1974).
3. Federation of American Scientists, Convention on the Prohibition of Military or Any Other Hostile Use of Environmental Modification Techniques (1980); www.fas.org/nuke/control/enmod/text/environ2.htm.

Geoengineering: Perilous Particles

THE POLICY FORUM "END THE DEADLOCK ON governance of geoengineering research" (E. A. Parson and D. W. Keith, 15 March, p. 1278) advances proposals for governmental regulation of geoengineering, the use of technologies to alter the climate in an attempt to mitigate the impacts of global warming. This strategy includes injecting metallic nanoparticles (e.g., aluminum or aluminum oxide) into the upper atmosphere to scatter incoming sunlight back to space. The authors suggest a moratorium on large-scale geoengineering but argue in favor of small-scale field research as well as model and laboratory studies. We believe that even small-scale tests are risky.

The notion of deliberate release of vast amounts of metallic nanoparticles into the atmosphere seems completely unsound in light of the known toxic potential of such materials (*1*). For instance, metal oxide nanoparticles trigger pulmonary inflammation in animal models (*2*). In addition, concerns have been raised about the consequences of engineered nanoparticles entering the environment, especially those that are nondegradable and persistent (*3*). In a recent study, aluminum nanoparticles were found to be toxic for freshwater algae (*4*).

Releasing nanoparticles into the sky to reduce global warming may only add insult to injury.

BENGT FADEEL,* HANNA L. KARLSSON, KUNAL BHATTACHARYA

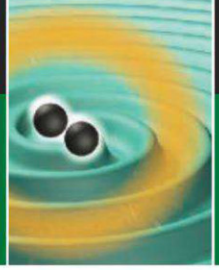
Division of Molecular Toxicology, Institute of Environmental Medicine, Karolinska Institutet, 171 77 Stockholm, Sweden.

*Corresponding author. E-mail: bengt.fadeel@ki.se

Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the past 3 months or matters of general interest. Letters are not acknowledged upon receipt. Whether published in full or in part, Letters are subject to editing for clarity and space. Letters submitted, published, or posted elsewhere, in print or online, will be disqualified. To submit a Letter, go to www.submit2science.org.

CREDIT: INGA SPENCE/ALAMY



A gold rush for gravitational waves

555

Seabirds as pollution monitors

556

References

1. A. Nel, T. Xia, L. Mädler, N. Li, *Science* **311**, 622 (2006).
2. W. S. Cho *et al.*, *Environ. Health Perspect.* **118**, 1699 (2010).
3. R. Behra, H. Krug, *Nat. Nanotechnol.* **3**, 253 (2008).
4. S. Pakrashi *et al.*, *Aquat. Toxicol.* **132–133C**, 34 (2013).

CORRECTIONS AND CLARIFICATIONS

News Focus: "When early hominins got a grip" by A. Gibbons (26 April, p. 426). The site of Kaitio near Lake Turkana was identified as being in Ethiopia. It is in Kenya. The HTML and PDF versions online have been corrected.

TECHNICAL COMMENT ABSTRACTS

Comment on "Bateman in Nature: Predation on Offspring Reduces the Potential for Sexual Selection"

Göran Arnqvist

Byers and Dunn (Reports, 9 November 2012, p. 802) reported that sexual selection and natural selection are closely related in a wild population of pronghorns. Here, I argue that this conclusion is incorrect. Their main finding is due to the fact that, unsurprisingly, juvenile mortality and juvenile survival are negatively related across years.

Full text at http://dx.doi.org/10.1126/science.1233413

Comment on "Bateman in Nature: Predation on Offspring Reduces the Potential for Sexual Selection"

P. Bergeron, A. M. Martin, D. Garant, F. Pelletier

Byers and Dunn (Reports, 9 November 2012, p. 802) claimed that predation on offspring reduced the potential for sexual selection in pronghorn. We argue that the potential for sexual selection is not affected by random offspring mortality when relative reproductive success is considered and increases when measured with the opportunity for selection, a metric that describes the potential for selection.

Full text at http://dx.doi.org/10.1126/science.1233246

Comment on "Bateman in Nature: Predation on Offspring Reduces the Potential for Sexual Selection"

Steven A. Ramm, Rudy M. Jonker, Klaus Reinhold, Tamás Székely, Fritz Trillmich, Tim Schmoll, Holger Schielzeth, Robert P. Freckleton

Byers and Dunn's (Reports, 9 November 2012, p. 802) conclusion that predation constrains sexual selection is problematic for three reasons: their nonstandard calculation of Bateman slopes; their assertion that random processes do not influence reproductive success; and the statistically unjustifiable use of 6 variables to explain just 10 observations.

Full text at http://dx.doi.org/10.1126/science.1233298

Response to Comments on "Bateman in Nature: Predation on Offspring Reduces the Potential for Sexual Selection"

John Byers and Stacey Dunn

Commenters objected to the way that we counted matings and offspring to calculate Bateman slopes and disagreed with our contention that predation on offspring can decrease the potential for sexual selection. We clarify what may have been misunderstandings to argue that our methods, analyses, and conclusions are correct.

Full text at http://dx.doi.org/10.1126/science.1233500

LIFE IN SCIENCE

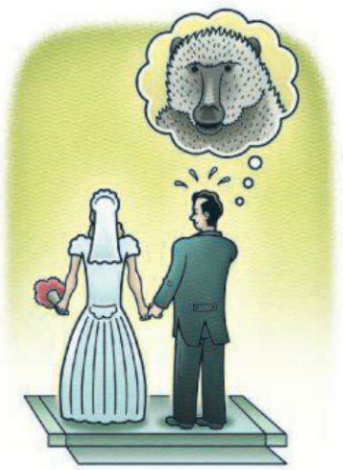
All in the Family

When I was in college, I spent a summer studying a troop of baboons living in a primate facility. Every day, I would record which baboons were socializing with one another. The dominant baboon was a 10-year-old male named Owen. Usually, Owen just hung out atop the 20-foot perch and threatened anyone who entered the observation area.

One day, my 17-year-old cousin (let's call him Brandon) came to visit. As we walked to the primate center, I explained that we would be in an observation space that separated us from the baboons by large, vertical bars placed a few inches apart. "Just remember," I told him, "baboons threaten each other by showing their teeth and staring each other in the face. So try not to do that."

When Brandon entered the observation room, a baby baboon reached through the bars and started tugging on Brandon's shirt, begging for treats. My cousin reached out his hand with raisins, looked the infant right in the face, and with a big smile exclaimed, "Look how cute he is!"

Immediately, Owen—who had been watching intently from his perch—jumped to the ground and came rushing toward Brandon. My cousin had unknowingly threatened the youngest member of Owen's troop, and Owen was going to teach him a lesson. Owen grabbed Brandon with his hands, pulling him up against the bars.



My cousin tried to push himself away and yelled for me to do something. I knew there was no way Owen could get out or really hurt Brandon, and at worse a piece of Brandon's shirt would probably tear off in Owen's hand, bringing this all to an end. Because baboons are troop animals and the other males were watching to see what I would do, I decided the best strategy was to patiently wait for Owen to finish with his threat. Finally, Owen let go of Brandon—who fell back from the bars onto the ground—and climbed back up to his perch, howling and jumping, declaring victory over my cousin. The whole troop was in chaos; the females and younger baboons ran into the enclosed section of the facility, and the other males were now threatening me through the bars. I put my stuff away and helped Brandon up from the ground.

"Now what do we do?" he asked.

"Now we go play tennis," I said as we left the observation room, "because you've pretty much ruined any data I might collect today."

Owen is now famous in our family. The story has often been retold at family reunions, and for years Brandon received occasional postcards from our other cousins (there are 21 of us) of baboons, with the poorly scrawled words "I am watching you. –Owen." During the rehearsal dinner for Brandon's wedding, the priest asked his brother if there were any good stories to share. The next day, during his homily, the priest told Brandon, "Owen is very sorry he is not here to wish you happiness." Brandon's fiancée whispered, "Who is Owen?" Brandon just shook his head and whispered, "Never mind."

GARTH FOWLER

Washington, DC 20002, USA. E-mail: garp@uw.edu

EDITOR'S NOTE

This is an occasional feature highlighting some of the day-to-day humorous realities that face our readers. Can you top this? Submit your best stories at www.submit2science.org.

CREDIT: PETER AND MARIA HOEY/WWW.PETERHOEY.COM

PSYCHOLOGY

Thinking, Broad and Deep

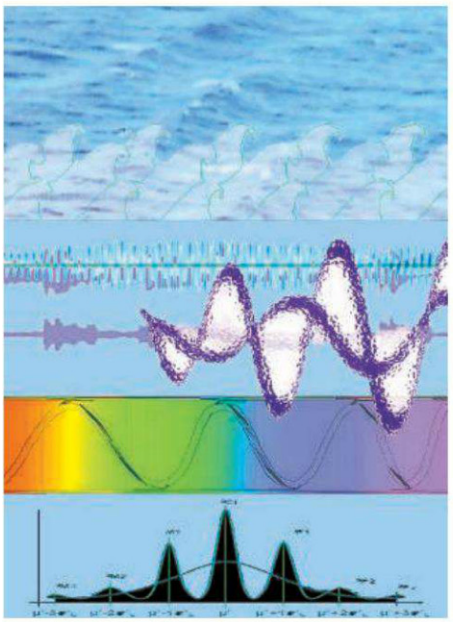
Keith J. Holyoak

What makes human thinking special? Addressing the American Psychological Association a half century ago, physicist Robert Oppenheimer made the case for the centrality of analogy: "Whether or not we talk of discovery or of invention, analogy is inevitable in human thought, because we come to new things in science with what equipment we have, which is how we have learned to think, and above all how we have learned to think about the relatedness of things" (*1*). In *Surfaces and Essences*, Douglas Hofstadter (Indiana University) and Emmanuel Sander [University of Paris (Saint-Denis)] build the case that the ability to see analogies indeed forms the core of human thinking—the way we "think about the relatedness of things."

Surfaces and Essences
Analogy as the Fuel and Fire of Thinking
by Douglas Hofstadter and Emmanuel Sander
Basic Books, New York, 2013. 592 pp. $35, C$38. ISBN 9780465018475.

No one is better equipped to make the case for analogy than the senior author. Hofstadter, analogist extraordinaire, burst onto the stage of cognitive science in 1979 with his Pulitzer Prize–winning *Gödel, Escher, Bach*, in which he created playful analogies and allegorical dialogues to illuminate such mathematical abstractions as recursion and undecidability (*2*). His intellectual breadth—trained in mathematics and physics, professor of computer science and psychology, artist, translator of Russian poetry and French novels—brings with it the capacity to see long-distance connections between situations and ideas, abstract their essences, and ground these abstractions in illuminating analogies. *Surfaces and Essences*, though not a sequel to *Gödel, Escher, Bach*, inherits a good deal of its intellectual focus and playful spirit. Like a strong marriage, the collaboration between Hofstadter and Sander, a French psychologist, is sufficiently seamless that the book reads as a single voice. Sander deserves credit for bringing in a salutary dose of psychological research, particularly on mathematical problem-solving and education, where goals and causal understanding are critical in distinguishing essence from surface. The mathematician soars among pure patterns; the psychologist stays rooted in human concerns. Their Anglo-Francophone collaboration reifies the art of translation, one of the most complex types of analogy-making. Rather than producing a conventional translation from source to target language, the authors worked in parallel on English and French versions. The book includes an illuminating self-referential sketch of its bilingual origin.

Hofstadter and Sander's thesis—analogy is the core of cognition—is less (or more) radical than it might sound, as they extend analogy to include "categorization through analogy-making." One situation is compared to another—two faces, two dogs, a heart and a pump—yielding a proto-category, to be refined by additional examples. Categories are not fixed and final products but are endlessly extensible by analogy. Waves on water come to embrace sound waves, then light waves, then spin waves, and then probability waves, as the concept wave becomes increasingly abstract. The authors convincingly refute those enthusiasts of embodied cognition who assume that because concepts are typically grounded in human perception and action, abstraction has been explained away. No: "abstraction is key, and to leave it out of one's theory of thinking is to miss the boat by a wide margin."



Abstracted by analogy. Waves on water are mapped to sound waves, then light waves, and then probability waves.

CREDIT: AIROM BLEICHER

The book grounds its abstractions in a garden of delightful examples: analogies based on words, phrases, metaphors, and proverbs; "me, too" stories where one person's anecdote elicits an analogical reminding in a listener; slips of action and of the tongue. Lofty scientific analogies are foreshadowed by the "banalogies" of everyday cognition. An elderly father driving by a cemetery baffles his adult son with the remark, "This is where all four of your grandkids were born"—the intended "all four of your grandparents are buried" fell victim to analogical slippage. As a child analogizes a toy truck to a real one, so Galileo analogized from Earth's one-of-a-kind Moon to hypothesize the moons of Jupiter (exemplifying "meta-analogy"). In the final chapter, Hofstadter the mathematician-physicist provides a compelling exposition of the analogical origins of number concepts and Einstein's relativity theory. Over a page of the book's index is devoted to the entry "lists" (e.g., "of abstract uses of 'mother,'" "of *sour grapes* situations," "of computer concepts used in daily life"). Extensive endnotes and references provide an excellent overview of scholarly sources.

The authors provide a cornucopia of analogical examples and qualitative insights but largely bypass computational and neural constraints on analogy (e.g., the critical concept of "binding" is not discussed). The influence of prior experience on cognition is indeed ubiquitous, but is it always "analogy"? People (and other animals) also learn by conditioning, passive accumulation of statistical associations, and other implicit mechanisms [e.g., (*3*)]. The authors appear to be of two minds on the question of whether analogy is unique to humans. Like Darwin before them (*4*), they are avowed dog fanciers and similarly apply the most naïve of analogies—anthropomorphism—to their canine friends: "categorization for a dog is clearly the creation of analogical bridges to prior knowledge." But a later section titled "What Makes *Homo Sapiens Sapiens* Sapiens?" reads like a retraction. The uniquely human core of analogy—the ability to encode and flexibly re-represent "the relatedness of things"—accounts for the fact that to date, scientists have emerged in only one species.

Surfaces and Essences warrants a place alongside *Gödel, Escher, Bach* and major recent treatments of human cognition (*5*). Analogy is not the endpoint of understanding, but its indispensable beginning. As Oppenheimer observed, "We cannot learn that we have made a mistake unless we can

The reviewer is at the Department of Psychology, University of California, Los Angeles, Los Angeles, CA 90095–1563, USA. E-mail: holyoak@lifesci.ucla.edu

make a mistake; and our mistake is almost always in the form of an analogy to some other piece of experience" (*1*).

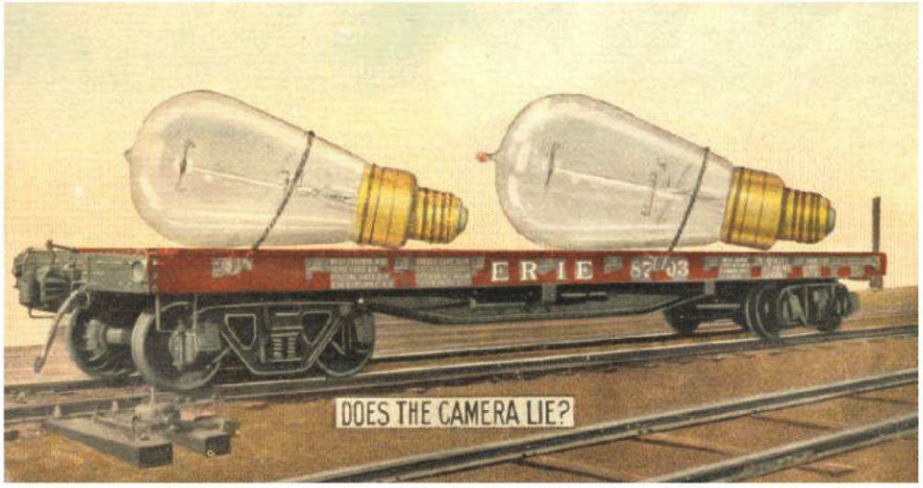
References

1. R. Oppenheimer, *Am. Psychol.* **11**, 127 (1956).
2. D. R. Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid* (Basic, New York, 1979).
3. B. J. Knowlton, J. A. Mangels, L. R. Squire, *Science* **273**, 1399 (1996).
4. D. C. Penn, K. J. Holyoak, D. J. Povinelli, *Behav. Brain Sci.* **31**, 109, discussion 130 (2008).
5. D. Kahneman, *Thinking, Fast and Slow* (Farrar, Straus and Giroux, New York, 2011); reviewed in (*6*).
6. S. J. Sherman, *Science* **334**, 1062 (2011).

10.1126/science.1236643

EXHIBITIONS: PHOTOGRAPHY

People Have Been Faking It for Years



Does the Camera Lie? (chromolithograph, unidentified American artist, circa 1910).

CREDIT: COURTESY THE METROPOLITAN MUSEUM OF ART

What's old is new again. Or maybe it never really became old in the first place. Recent press coverage has spotlighted covertly digitally manipulated photographs of war-torn areas and even instances to which a publisher freely admits, such as *The Washington Post*'s 13 January 2012 front-page, composite-exposure photo (*1*). Such events capture the public's attention and may lead one to believe the practice of photo manipulation is a relatively recent phenomenon, existing in great part because it is possible with today's photographic technology. As the enlightening exhibition *Faking It: Manipulated Photography Before Photoshop* reminds us, today's news items are only the latest examples of a practice that is nearly as old as the medium itself.

The exhibition—organized by the Metropolitan Museum of Art and curator Mia Fineman, sponsored by Adobe (creator of Photoshop), and currently on display at the National Gallery of Art in Washington, DC—explores manipulation of nondigital photography from the 1840s through the early 1990s. The images on display "were modified in the service of art, politics, journalism, entertainment, or commerce."

Whereas nowadays we think of photographic manipulation as being done because we can, it was originally performed to compensate for what photography couldn't achieve. Early photography had some rather severe technical limitations, and early photographers developed workarounds that would create a more realistic representation of the photographic subject, a more aesthetically pleasing image, or both. As the medium progressed and evolved, so did the techniques for manipulation and the intentions of photographers. There were still those for whom aesthetics and realism remained a key concern, but others delved into photography as a purely artistic medium, exploring surrealistic and fanciful imagery. Sometimes the manipulation was obvious, and timelessly delightful, as in the tall-tale Americana postcards of the early 20th century. In other cases, photographers seemed to want their doctoring to serve a seamless illusion. And for whatever reason, images of people with decapitated heads were apparently quite the rage.

Photography and politics are natural bedfellows, and photo manipulation has long been part of the political arena. Manipulated images have been used in the service of promoting particular ideologies and also to caricature politics. The most iconic portrait of Chairman Mao was in fact highly retouched by Chen Shilin, "to project an image of flawless benevolence," while during the Vietnam War, Wegee turned President Johnson into a latter-day Pinocchio.

One particularly interesting concept presented in the exhibition is the public's own responsibility for photo manipulation in journalism. Early on, newspapers and magazines relied on illustrations for visual accompaniment to a story. As photography gained more prominence, the public came to expect and even demand it instead. However, many stories simply couldn't be photographed. That inconvenience did not stop publishers, who turned to creating photos to support stories—either by altering existing shots of certain subjects or even fabricating scenes entirely. The exhibition and accompanying catalog provide several examples of such manipulations, although these do not establish whether it is fair to blame the public for publishers' bowing to the pressure for more photographic representation.

Some take exception to the adage "the camera doesn't lie." But the camera is a tool—a machine incapable of such a sentient act as lying. It can only do what the photographer sets it to do. Of course, photographers may lie using the tools at their disposal. And photography is by its very nature a manipulated task. There are infinite ways of manipulating a photographic image—by selecting a different f-stop when first depressing the shutter, for example, or varying the development of an image in the darkroom—that are not only valid but necessary. The ways that photographers have taken these tools and expanded upon them, for purposes well intentioned or malicious, lighthearted or grave, present a deep well from which *Faking It* draws riches.

–Yael Fitzpatrick

Faking It
Manipulated Photography Before Photoshop

Mia Fineman, curator
Metropolitan Museum of Art, New York, through 27 January 2013. National Gallery of Art, Washington, DC, through 5 May. Museum of Fine Arts, Houston, 2 June to 25 August 2013. www.nga.gov/exhibitions/faking.shtm

Faking It
Manipulated Photography Before Photoshop

by Mia Fineman
Metropolitan Museum of Art, New York, 2012. 296 pp. $60. ISBN 9780300185010.

References

1. www.washingtonpost.com/blogs/ask-the-post/post/post-photography-and-the-use of-high-dynamic-range/2012/01/17/gIQAHGIg5P_blog.html.

10.1126/science.1238524

RESEARCH ETHICS

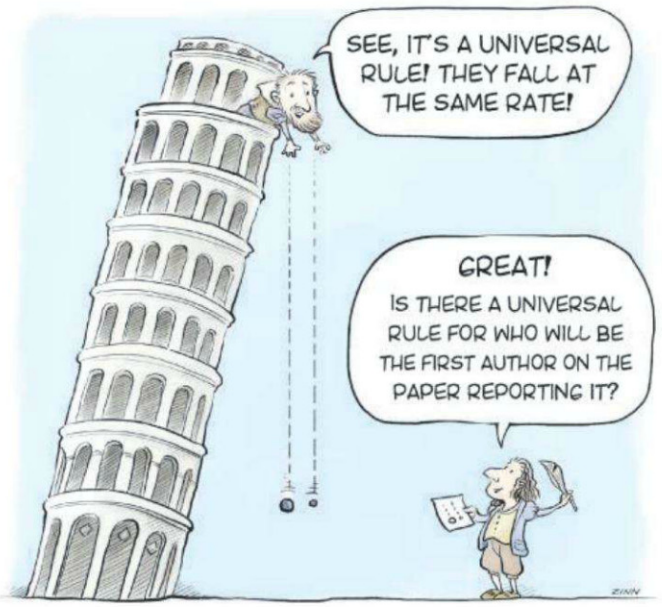
Global Research Integrity Training

Globalization of responsible conduct of research training should be grounded on shared professional standards.

Nicholas H. Steneck

Misconduct in research remains an unresolved problem despite over 30 years of local, national, and global efforts to reduce or eliminate it. For every confirmed case of serious misbehavior or misconduct, 10 or more may exist along with many more instances of lesser unprofessional practice (*1*, *2*). Beneath the surface of visible cases, the underlying structure supporting the responsible conduct of research (RCR) has significant flaws (*3*).

In May 2013, the Third World Conference on Research Integrity (WCRI) and the meeting of the Global Research Council (GRC) will offer opportunities to improve the reliability and accountability of the global response to misconduct in research. I discuss how globalizing RCR training can contribute to these new, long-overdue worldwide efforts.



Researchers and policy-makers are working to improve responses to misconduct in research. Australia, Brazil, Canada, China, and Ireland have taken recent steps to develop or strengthen misconduct policies. Universities are taking a more active role in promoting integrity, illustrated by the recent Universities UK *Concordat* (*4*) and meetings held by the League of European Research Universities (*5*) and Universitas 21 (*6*). The InterAcademy Council recently issued a report and code of conduct (*7*), providing another option to the global *Singapore Statement on Research Integrity* developed at the Second WCRI (*8*). Still, we need coordination of global efforts to promote integrity through RCR training.

Promoting Integrity Through Training

Misconduct and misbehaviors in research are products of decisions individual researchers make. Policy-makers assumed that training researchers could reduce these problems. In 1989, the U.S. National Institutes of Health (NIH) and the Alcohol, Drug Abuse, and Mental Health Administration issued the first U.S. RCR training requirement to test informally whether more emphasis on training would make a difference (*9*).

The new requirement had a substantial impact. Over the next decade, new training courses and textbooks surfaced. The U.S. Office of Research Integrity (ORI) initiated RCR programs to broaden coverage and develop research on research integrity. More recently, the U.S. National Science Foundation (NSF) funded a national research ethics center, the Ethics CORE (Collaborative Online Resource Environment) (*10*). These "carrots" along with the regulatory "sticks" spurred U.S. development of RCR (*11*). Improved RCR training is now variously supported in other countries as well.

Despite these encouraging developments, the future of RCR training is far from certain. Studies of effectiveness have mostly been inconclusive. Students respond favorably and may learn skills such as rational problem-solving (*12*, *13*), but long-term impacts on behavior have not been demonstrated (*14*). Weak financial support from governments and other funders places the burden for training on research institutions. The lack of evidence for benefits combined with the burden for delivery will make it difficult for the globalization of RCR training to continue. Are there reasons and strategies for continuing the globalization of research integrity training in an increasingly demanding world?

Globalizing RCR Training

Calling for development of the Global Research Council, former NSF Director Subra Suresh wrote: "The most fundamental barriers to bilateral and multilateral international collaborations are disparate standards for scientific merit review and differences in the infrastructures that ensure professional ethics and scientific integrity" (*15*). Shared acceptance of the "rules" for ethics and integrity are as essential to collaboration and progress in research as agreement on the basic laws of nature.

Globalization of RCR training would harmonize and gain greater support for the common rules and professional standards for responsible research. However, to fulfill this role, RCR training first needs itself to be harmonized.

How Can Globalization Be Accomplished?

Common standards. If basic framework development is left in the hands of individual countries and then delegated to institutions, departments, and instructors, the problem of disparate standards will continue, or worsen. Common professional standards must be globally defined and adopted. The regional and international codes mentioned above provide frameworks for developing training programs. A working group at the Second WCRI outlined a plan and proposed steps for harmonization of RCR training efforts (*16*). RCR training and common standards are on the agenda of this month's WCRI and GRC meetings. Governments, institutions, and researchers must recognize that RCR training must be planned as a common responsibility.

Common content. The easiest and most economical way to harmonize RCR training is to use common materials. There are differences in national regulations and discipline-specific professional practices. One course, textbook, or Web site will never serve all purposes, and there will always be a need for advanced, country- and discipline-specific

Director, Research Ethics and Integrity Program, Michigan Institute for Clinical and Health Research, University of Michigan, Ann Arbor, MI 48103, USA.

E-mail: nsteneck@umich.edu

CREDIT: DAVID ZINN/WWW.ZINNART.COM

training, delivered by researchers or mentors in small groups or one-on-one. But it is inefficient to have every country, institution, and instructor design courses on, for example, responsible publication practices, management of international collaborations, or the responsibility to report misconduct.

Considerable effort has gone into development of training materials over the past two-plus decades. Ethics CORE houses >5000 RCR training items (*10*). So many resources may be a boon to the instructor but a potential disaster for globalization of training. The pressing need is not more resources but better ways to categorize and evaluate content.

Collaborative training. Global harmonization of RCR training can be promoted and made more efficient via collaboration. This is most easily accomplished via the Web, as illustrated by the pioneering Collaborative Institutional Training Initiative (CITI). CITI's founders wanted to develop a professionally responsible training program for human subjects protections. Working across institutions, they developed an online system now used at over 2000 sites in more than 40 countries.

Collaborative online training has critics (*17*); its full potential has yet to be realized. Universitas 21 is using it to deliver case-based training to graduate schools. The Ethics CORE RCR modules provide a way of managing online training along with access to the full Ethics CORE library. Epigeum's Research Integrity courses, which this author organized and cowrote, aim to extend learning and interaction into peer communities and research settings.

Global adoption of common goals, content, and delivery mechanisms for RCR would raise awareness of responsible research practices on a uniform basis. Regardless of how they ultimately behave, all researchers should be provided with basic tools for acting responsibly. At present, they are not. Basic training is not available in many countries and unevenly taught even in the United States, where it is required of most researchers.

Redesigning the Experiment

Improved and expanded global RCR training will not necessarily reduce misconduct or improve integrity in research. A basic global strategy must include all factors that influence behavior once awareness has been raised. The initial RCR training experiment needs to be redesigned, focusing on four challenges:

Lack of clear goals. Planning and assessment are hindered by a lack of clear goals. The NSF asks only for assurance that training has been provided, but sets no standards for content or outcomes (*18*). The UK Concordat promotes "awareness among researchers of the standards and behaviours that are expected of them" (*4*). The NIH expects training to lead to the "application of established professional norms and ethical principles in the performance of all activities related to scientific research" (*19*). A survey of RCR instructors identified 50 distinct goals they were trying to achieve (*20*). Assessment of success or failure of RCR training must be based on clear, measurable goals.

Pedagogical disagreements. There are unresolved pedagogical disagreements that grow out of assumptions about influences on and ways to change research behavior. In developing training, emphasis is variously placed on moral reasoning and decision-making (*21*), climate (*22*), or the lack of proper understanding of professional norms (*23*). NIH suggests that RCR training should be integrated into research training, relevant to career stage, and include face-to-face opportunities with mentors and experienced researchers (*19*). The European Commission and the NSF see RCR training as a component of professional ethics; the NIH and countries such as Canada and Australia place RCR training in the context of regulation and professional responsibility. Strategies and approaches must be compared, particularly as the RCR experiment is expanded globally.

Research integrity metrics. Measuring success or failure requires common metrics. Researchers have developed tools to assess the prevalence of misbehavior (*1*), plagiarism and duplicate publication (*24*), the influence of climate (*25*), and ethical reasoning (*12*, *13*). There is no easy way to share and compare data. We lack the basic tools to conduct the rigorous experiments to compare approaches and to measure success.

Cost-benefit analysis. Cost is a factor in planning RCR training. In-person instruction may seem like the best approach, but enhanced, Web-based instruction may be able to accomplish some of the same goals, with significant savings in cost and time.

RCR training could provide potential savings. One less misconduct case could save a university $10,000 to $100,000 in investigation costs (*26*). One less serial plagiarizer could save editors the cost of retracting 50 to 150+ articles (*27*). More rigorous and honest research design could reduce inconclusive outcomes in clinical trials (*28*). Better training in regulatory responsibility could make human subjects research proposal review more efficient and reduce financial risks of noncompliance (*29*). Such savings should be weighed against the small amount most institutions spend promoting RCR through training.

It is time to use the same critical approach to RCR training as used for any research problem—identify the problem, develop hypotheses about causes, run experiments, interpret the results, reach conclusions—to figure out whether improved training is a reasonable way to improve integrity in research and, if so, which approaches to training are both efficient and effective.

References and Notes

1. B. C. Martinson, M. S. Anderson, R. de Vries, *Nature* **435**, 737 (2005).
2. D. Fanelli, *PLoS ONE* **4**, e5738 (2009).
3. N. H. Steneck, *Sci. Eng. Ethics* **12**, 53 (2006).
4. Universities U. K., *The Concordat to Support Research Integrity* (Universities UK, London, 2012).
5. League of European Research Universities (2012); www.leru.org/index.php/public/calendar/research-integrity-forum/.
6. Universitas 21, "Research integrity workshop leads to new collaboration," *Universitas 21 News*, 9 January 2013.
7. InterAcademy Council and IAP, *Responsible Conduct in the Global Research Enterprise: A Policy Report* (InterAcademy Council, Amsterdam, 2012).
8. Second World Conference on Research Integrity, Singapore, 21 to 24 July 2010 (World Conferences on Research Integrity, 2010).
9. Alcohol, Drug Abuse, and Mental Health Administration, NIH. *NIH Guide Grants Contracts* **18**, 1 (1989).
10. Ethics CORE (University of Illinois); https://nationalethics-center.org/.
11. N. H. Steneck, R. E. Bulger, *Acad. Med.* **82**, 829 (2007).
12. M. D. Mumford *et al.*, *Ethics Behav.* **17**, 337 (2007).
13. M. D. Mumford *et al.*, *Ethics Behav.* **18**, 315 (2008).
14. M. S. Anderson *et al.*, *Acad. Med.* **82**, 853 (2007).
15. S. Suresh, *Science* **333**, 802 (2011).
16. N. H. Steneck, M. Kalichman, M. Axelsen, in *Promoting Research Integrity in a Global Environment*, T. Mayer and N. H. Steneck, Eds. (World Scientific, Singapore, 2012).
17. J. J. Freyd, *J. Trauma Dissoc.* **12**, 475 (2011).
18. National Science Foundation, *Fed. Regist.* **74**, 42126 (2009).
19. Office of the Director, NIH, Update on the requirement for instruction in the responsible conduct of research (NOT-OD-10-019, NIH, Bethesda, MD, 2011).
20. M. W. Kalichman, D. K. Plemmons, *Acad. Med.* **82**, 846 (2007).
21. A. L. Antes *et al.*, *J. Empir. Res. Hum. Res. Ethics* **2**, 15 (2007).
22. B. C. Martinson, A. L. Crain, R. De Vries, M. S. Anderson, *J. Empir. Res. Hum. Res. Ethics* **5**, 67 (2010).
23. E. Heitman, R. E. Bulger, *AIR* **12**, 207 (2005).
24. M. Errami, Z. Sun, T. C. Long, A. C. George, H. R. Garner, *Nucleic Acids Res.* **37**, (Database), D921 (2009).
25. B. C. Martinson, C. R. Thrush, A. L. Crain, *Sci. Eng. Ethics* (2012). 10.1007/s11948-012-9410-7
26. A. M. Michalek, A. D. Hutson, C. P. Wicher, D. L. Trump, *PLoS Med.* **7**, e1000318 (2010).
27. S. Stretton *et al.*, *Curr. Med. Res. Opin.* **28**, 1575 (2012).
28. B. Djulbegovic *et al.*, *Arch. Intern. Med.* **168**, 632 (2008).
29. P. Keith-Spiegel, G. P. Koocher, *Ethics Behav.* **15**, 339 (2005).

Acknowledgment: The author is a consultant, adviser, and author for Epigeum online RCR training courses. His RCR training activities at the University of Michigan are supported by NIH grant 2UL1TR000433.

10.1126/science.1236373

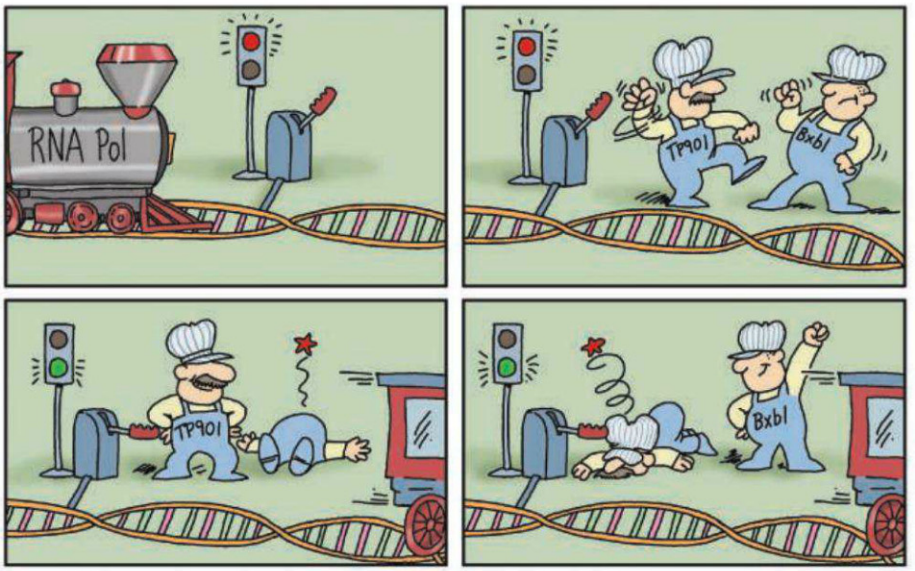
ENGINEERING

Recombinatorial Logic

The use of sequence inversions allows all 16 possible logic gates for two inputs to be realized in bacterial cells.

Yaakov Benenson

Logic gates evoke images of circuit boards, but cells are arguably equally good in relying on logic computations. A classic example is the Lac operon, which activates itself upon the condition "lactose AND NOT glucose" (*1*). In recent years, there have been multiple reports on rationally designed, genetically encoded logic gates and circuits in living cells (*2*). Just like the Lac operon, these gates receive two or more molecular signals (inputs) and generate a product (output) whose level is logically linked to the inputs. Sixteen different logic connections are possible with two inputs and one output, but many of these operations have remained refractory to rational design. The trickiest of these gates usually make general statements about the inputs without referring to their exact values, such as "both inputs are the same" (an XNOR gate) or "two inputs are different" (an XOR gate). Two studies, one on page 599 of this issue by Bonnet *et al*. (*3*) and one by Siuti *et al*. (*4*), describe approaches that produce any of the 16 gates, including the notorious XNOR and XOR, in a compact manner by making relatively minor tweaks to the gates' genetic building blocks.



Controlling the signal. Transcription in a bacterial cell can be controlled by engineered logic gates. The approach taken by Bonnet *et al*. to create an XOR gate is illustrated by a cartoon that depicts RNA polymerase as a locomotive awaiting its go-ahead signal. A terminator sequence prevents transcription in two cases of activity by two recombinase enzymes, TP901 and Bxb1 if neither is active (**top left**), or if both TP901 and Bxb1 are active, effectively neutralizing each other's action (**top right**). When there is exactly one recombinase active (**bottom**), the go-ahead signal is given by inverting the terminator and allowing RNA to be synthesized.

The uneven progress in implementing synthetic gates has been mirrored by a recent argument that the natural evolution of some regulatory logic gates is much more difficult than for others (*5*). Indeed, the success of the approaches of Bonnet *et al*. and Siuti *et al*. resulted from repurposing phage-derived molecular tools whose natural role is not to control gene expression, but rather to insert phage DNA into host genomes: site-specific serine recombinases, or integrases. An integrase recognizes a pair of specific, nonidentical sequences called attP (on the phage) and attB (on the host) that, after recombination and phage insertion, are converted into a pair of sites called attL and attR that no longer serve as recombinase substrates. By intentionally framing an engineered DNA sequence by attP and attB sites, it is possible to trick the recombinase into inverting this sequence irreversibly. The "irreversibility" is not the whole story; an integrase can recognize attL and attR sequences in the presence of another protein called recombination directionality factor (RDF).

Department of Biosystems Science and Engineering, Swiss Federal Institute of Technology (ETH Zürich), Mattenstrasse 26, 4058 Basel, Switzerland. E-mail: kobi.benenson@bsse.ethz.ch

Using recombinases to control gene expression is not unheard of in nature: fimB/E recombinases flip a promoter controlling *Escherichia coli* virulence (*6*). In genetic engineering (*7*), recombinases have been used for insertion, excision, or inversion of DNA sequences to create either well-formed expression units or mutants incapable of gene expression. The devices in the recent reports deal exclusively in inversions because the underlying biochemical process is highly efficient and because it can be selectively reversed with the help of RDFs. The major advance comes from controlling the same gene with more than one recombinase, requiring certain activity patterns of multiple enzymes to create transcriptionally active configurations. This approach, together with ingenious design of the specific recombinase substrates, enabled the long-sought diversity of logical operations.

The gates built by Siuti *et al*. flipped gene promoters, a transcriptional terminator, and the gene-coding sequence. Bonnet *et al*. relied mostly on terminators and occasionally on promoters. For example, they implemented the XOR gate by placing a terminator in front of the gene-coding sequence and then flanking the terminator with two nested pairs of attP/attB sites for two different integrases. If either one of the enzymes was active, the terminator was flipped backward and out of context, allowing transcription to proceed. If both were present, the terminator would be inverted once and then again back into its blocking position (see the figure). The former study used a similar nested architecture but with a promoter instead of a terminator: A backward-facing promoter was restored into the correct orientation by either integrase, but was flipped back by the second one.

The irreversible nature of DNA modification sets recombinase-based gates apart from earlier work, where reversible input-output interactions caused the output to track changing inputs dynamically. In the new devices, integrase inputs perform irreversible modifications to the output DNA, so the logic conditions become more complex than simple "here and now." Thus, an AND gate generates the output if both inputs had been active in the gate's past, not necessarily at the same time.

CREDIT: JOE SUTLIFF/WWW.CDAD.COM/JOE

History-dependent operation invokes closely related concepts of sequential logic and finite-state machines. The latter can be described as a memory cell with an input receiver. The cell can store one piece of information (state) that changes when an input arrives. A number of studies explicitly explored recombinases as the tool for state machines and counters (*8*, *9*), and a rewritable memory unit had been implemented earlier using a bidirectional recombination process (*10*).

The integrase approach enables truly digital logic. In earlier work, a gene controlled by reversible inputs such as transcription factors (*11*, *12*) or microRNA (*13*) could be fully induced or repressed but could also exhibit intermediate expression levels. A single integrase-controlled gene can exist only in a finite number of states, and unless explicitly intended otherwise (*4*), it can be either fully active or completely inactive. The reported experiments were performed with multiple-copy plasmids, and the output level in each bacterium reflected intracellular ratios of different plasmid states, resulting in large output heterogeneity. However, each individual plasmid could only exist in a fully on or off state—a result that could be confirmed through careful single-cell measurements with single-copy plasmids or stable chromosomal integration.

How exactly the integrase affects the output on a single-cell, single-gene level is an open question. In both reports the integrases were controlled via inducible promoters, and simultaneous observation of the inputs (e.g., through fluorescent fusion) and the output in individual cells has not been made. These measurements are important because in principle, even a very small amount of integrase, given enough time, could induce switching and make the gates overly sensitive. Bonnet *et al*. made important first steps toward such characterization. They quantified integrase indirectly by substituting it with green fluorescent protein and inferred a population-averaged link between the integrase and the output. They observed signal attenuation and amplification at low and intermediate integrase levels, respectively. Both are highly desirable features in digital signal processing.

Recombinase gates represent a new paradigm in synthetic circuit design. The integration of logic, memory, input reset, and digital states makes the experimental systems reported by Bonnet *et al*. and Siuti *et al*. interesting test cases for the theory of biological networks. These gates are perhaps the farthest away from any natural counterpart and are thus likely to occupy synthetic and systems biologists in the years to come.

References

1. F. Jacob, J. Monod, *J. Mol. Biol.* **3**, 318 (1961).
2. Y. Benenson, *Nat. Rev. Genet.* **13**, 455 (2012).
3. J. Bonnet, P. Yin, M. E. Ortiz, P. Subsoontorn, D. Endy, *Science* **340**, 599 (2013); 10.1126/science.1232758.
4. P. Siuti, J. Yazbek, T. K. Lu, *Nat. Biotechnol.* 10.1038/nbt.2510 (2013).
5. L. G. Valiant, *J. ACM* **56**, 3 (2009).
6. M. S. McClain, I. C. Blomfield, B. I. Eisenstein, *J. Bacteriol.* **173**, 5308 (1991).
7. E. C. Dale, D. W. Ow, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 10558 (1991).
8. T. S. Ham, S. K. Lee, J. D. Keasling, A. P. Arkin, *PLoS ONE* **3**, e2815 (2008).
9. A. E. Friedland *et al.*, *Science* **324**, 1199 (2009).
10. J. Bonnet, P. Subsoontorn, D. Endy, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8884 (2012).
11. A. Tamsir, J. J. Tabor, C. A. Voigt, *Nature* **469**, 212 (2011).
12. S. Ausländer, D. Ausländer, M. Müller, M. Wieland, M. Fussenegger, *Nature* **487**, 123 (2012).
13. Z. Xie, L. Wroblewska, L. Prochazka, R. Weiss, Y. Benenson, *Science* **333**, 1307 (2011).

10.1126/science.1237738

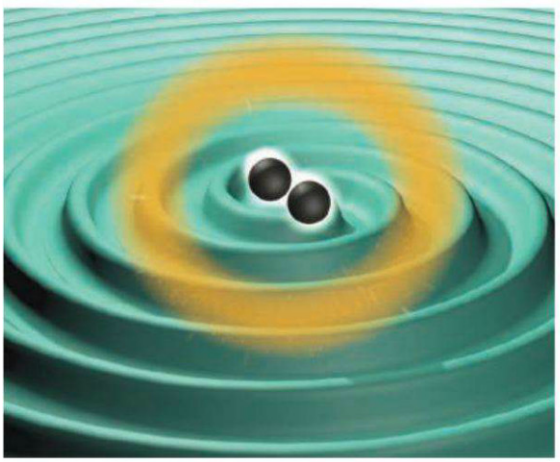
ASTROPHYSICS

Seeing Gravitational Waves

A suite of observatories will be needed to detect the visible electromagnetic counterpart of gravitational waves.

Mansi M. Kasliwal

Gravity is responsible for the long-range order of the universe. Using Einstein's general relativity, we now think of gravity as the geometrical curvature of the four-dimensional fabric of space-time (*1*). Extreme cosmological events such as the merging of neutron stars or black holes induce ripples in the fabric of space-time (see the figure). However, these ripples, or gravitational waves, are extremely weak, and their detection has remained elusive. To measure the small signal, an interferometric detector is required that can detect strain to one part in 10^{21} (that is, a billionth of a nanometer for a kilometer-length interferometer). Such extreme gravity events are also rare, occurring only once every 10,000 years per galaxy (*2*). An advanced version of such a detector is designed to find gravitational waves on a regular basis (roughly tens of events annually) beginning in 2017 (*3*). This heroic experiment alone will be somewhat unsatisfying—gravitational wave interferometers will only be able to hear the wave and detect when something happens (literally "hear" as the operational frequency of tens to thousands of Hertz overlaps with the human auditory range). The interferometers will be blind to exactly where the merger occurs. To locate the source of the gravitational waves, collaboration between the physics and the astronomy communities together with extensive simulations are under way (*4*).



Gravity wave detection. Merging neutron stars or black holes induce ripples in the fabric of space-time. The decade ahead promises to witness the first gravitational wave detections: to "hear" the sound waves with advanced gravitational wave interferometers and "see" the gold halo (produced via nucleosynthesis) with a slew of panchromatic telescopes.

There is a lot of activity in three complementary camps: instrumentalists, theorists, and observers. Instrumentalists are striving to set up a global network of advanced interferometers to localize the gravitational wave signals: the longer the baselines, the tighter the triangulation. The first such triangle to come online will be the Laser Interferometer Gravitational Wave Observatory at Hanford (LIGO-Hanford) (*5*), LIGO-Louisiana (*5*), and VIRGO-Italy

The Observatories, Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA. E-mail: mansi@obs.carnegiescience.edu

(*6*). Next will be LIGO-India (*7*) and Kagra-Japan (*8*). Median error regions of 50 (or 6) degrees^2 can be achieved with three (or five) interferometers (*4*, *9*). The network will be sensitive to mergers out to 400 Mpc (or 750 Mpc) for three (or five) interferometers (*2*, *4*). Low-latency gravitational wave localization volumes can be constructed within minutes to alert the astronomers to search for the relevant electromagnetic counterpart (*4*).

Theorists are working out the expected electromagnetic signature from neutron star and black hole mergers; the luminosity, time scale, and spectral energy distribution. The predicted counterpart is expected to be fainter than a supernova (but brighter than a nova), last for a few hours to a few days, and have red colors (*10*–*12*). Nucleosynthesis in the neutron-rich ejecta is expected to give elements with mass numbers ranging between 120 and 200. Indeed, the majority of gold in the universe may be produced during neutron star or black hole mergers (*12*, *13*).

Observational astronomers are mobilizing large telescopes across the entire electromagnetic spectrum. For the tiny fraction of jets (<2.5%) beamed toward us, an all-sky gamma-ray monitor (such as the Fermi and Swift space-based telescopes) detecting contemporaneous emission would be the most straightforward identification. Radio astronomers have also recently brought online a wide array of low-frequency antenna arrays to look for a contemporaneous pulse and upgraded existing facilities (e.g., Jansky Very Large Array) to improve mapping speed.

Optical astronomers have the strongest arsenal to search wide areas efficiently to look for a counterpart to all gravitational wave events. Very wide-field cameras on all sizes of telescopes are being built: Zwicky Transient Facility (35 degrees^2, 1.2 m; 2015), Dark Energy Camera (3 degrees^2, 4 m; 2012), and HyperSuprimeCam (1.8 degrees^2, 8.2 m; 2012). Opacity calculations suggest that there may be an emission peak in the infrared (*12*). Unfortunately, the infrared sky doesn't have wide-field instrumentation at this time. Astronomers have proposed to build the Synoptic All-Sky Infrared telescope (SASIR; 0.2 to 1 degrees^2) on the ground and the Wide-Field Infrared Survey Telescope (WFIRST; 0.3 degrees^2) in space.

The biggest challenge ahead is that the transient sky is extremely dynamic, which results in a large number of false positives (*4*, *14*). Because of the small solid angle occupied by galaxies on the sky, spatial coincidence with nearby galaxies can reduce the false positives from hundreds to just a few (*4*, *14*). Unfortunately, we do not know the location of half of the galaxies within the relevant horizon. Efforts are under way to complete our galaxy catalog using H-α narrow-band imaging in the optical and HI imaging in the radio. Systematic surveys are characterizing all types of transient phenomena in the local universe and have recently found multiple new distinct classes of elusive transients (*15*).

A complete inventory of transients and a complete catalog of nearby galaxies will empower us to find these needles in the haystack. Thus, there is a surge of excitement as the era of routine gravitational wave detection draws near—this search may prove to be the 21st-century gold rush.

References and Notes

1. C. W. Misner, K. S. Thorne, J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973).
2. J. Abadie *et al.*, *Class. Quantum Gravity* **27**, 173001 (2010).
3. J. Aasi *et al.*, LIGO technical document P1200087 (2013), https://dcc.ligo.org/LIGO-P1200087-v18/public.
4. S. Nissanke, M. Kasliwal, A. Georgieva, *Astrophys. J.* **767**, 124 (2013).
5. G. M. Harry *et al.*, *Class. Quantum Gravity* **27**, 084006 (2010).
6. T. Accadia *et al.*, *Class. Quantum Gravity* **28**, 114002 (2011).
7. B. Iyer *et al.*, LIGO technical document M1100296 (2011), https://dcc.ligo.org/LIGO-M1100296-v2/public.
8. C. Somiya, *Class. Quantum Gravity* **29**, 124007 (2012).
9. C. Schutz, *Class. Quantum Gravity* **28**, 125023 (2011).
10. X-L. Li, B. Paczyński, *Astrophys. J.* **507**, L59 (1998).
11. S. R. Kulkarni, http://arxiv.org/abs/astroph/0510256 (2005).
12. J. Barnes, D. Kasen, http://arxiv.org/abs/1303.5787 (2013).
13. B. D. Metzger *et al.*, *Mon. Not. R. Astron. Soc.* **406**, 2650 (2010).
14. S. R. Kulkarni, M. M. Kasliwal, http://arxiv.org/abs/0903.0218 (2009).
15. M. M. Kasliwal, *Publ. Astron. Soc. Aust.* **29**, 482 (2012).

Acknowledgments: The author acknowledges generous support from the Hubble Fellowship and the Carnegie-Princeton Fellowship.

10.1126/science.1235956

ENVIRONMENTAL SCIENCE

Tracking Marine Pollution

Seabird monitoring studies are providing a global picture of an increasing range of marine pollutants.

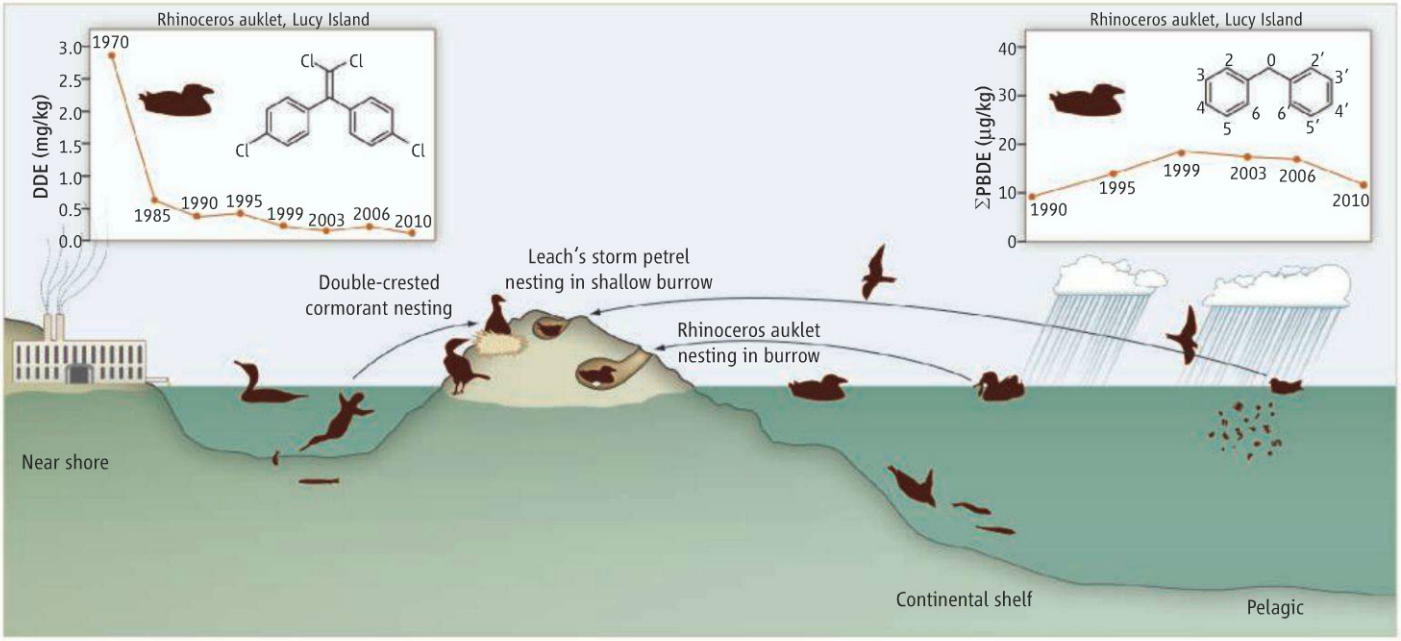
John E. Elliott[1] and Kyle H. Elliott[2]

Visit a beach almost anywhere and you will see plastic waste floating in the water and heaped above the tide lines. That debris is both a source and an overt signal of the even more pervasive contamination of marine biota by persistent chemicals. Present at ultra-trace levels but often highly toxic, chemical pollutants can be challenging to measure and understand. As the most problematic compounds biomagnify in food chains, sampling of marine top predators yields a global picture of ocean pollution.

[1]Science & Technology Branch, Environment Canada, Delta, BC V4K 3N2, Canada. [2]Department of Biological Sciences, University of Manitoba, Winnipeg, MB R3T 2N2, Canada. E-mail: john.elliott@ec.gc.ca

We have come a long way since Rachel Carson's *Silent Spring* and subsequent reports of chemical contamination of even the most remote ecosystems, including effects on the survival and reproduction of coastal seabirds (*1*). Use of the worst chlorinated chemicals was banned or severely restricted through the 2001 Stockholm Convention. Increasing mercury levels in marine wildlife, among other reasons, led to the adoption in January 2013 of the global Minamata Convention on Mercury. But with thousands of new chemicals introduced annually, some will inevitably slip through the regulatory net to become the next persistent global contaminants. There is thus a continuing need for efficient monitoring.

Mammals, particularly pinnipeds and cetaceans, are useful sentinels for marine pollution (*2*). However, seabirds have several practical advantages. Variance in organochlorine concentrations in seabirds is less than in fish or marine mammals; thus, a small sample size of seabirds can provide greater statistical power in addition to reduced environmental impact and cost of sampling (*3*). Seabirds range widely across oceans, feeding as they move, yet return annually to breed in a single central place (see the figure). In one afternoon at a seabird colony, a biologist can sample an area of ocean that would cost millions of dollars to investigate using a scientific vessel. Nonlethal samples of blood, feathers, oils, and biopsies each provide information on dif-



How seabirds sample the marine environment. Cormorants forage mainly on fish in near-shore environments. Auks, such as the rhinoceros auklet, feed on smaller fish and zooplankton on the continental shelf. Pelagic seabirds such as the Leach's storm petrel range across the offshore environment, feeding on zooplankton and larval fishes. They all return to breed in colonies, where they can be sampled for contaminants and tagged for tracking. As an example of the information that can be obtained, the graphs show data for two contaminants in rhinoceros auklet eggs at Lucy Island, northwest Canada: DDE, dichlorodiphenyldichloroethylene (a DDT metabolite), and PBDEs, polybrominated diphenyl ethers (a flame retardant).

ferent spatial and time scales (*4–7*). The lipid-rich avian egg, in particular, provides a simple matrix to measure the lipophilic chemicals that are of most concern (*8–14*).

Many studies provide a snapshot of contaminants in marine birds from locations around the globe. Some developed countries have also maintained consistent sampling efforts (*8–14*). For example, Environment Canada has monitored near-shore and offshore habitats of three oceans and the Great Lakes since the early 1970s (*8*, *9*, *14*). Other countries have devised similar schemes that vary in choice of seabirds and scope (*7*, *10–13*). Museum specimens have been used to examine trends from a century ago or longer—for example, affirming the rise in mercury levels associated with fossil fuel use in the 20th century (*5*).

Seabird egg monitoring documented the decline of the legacy persistent organic pollutants (POPs) such as dichlorodiphenyltrichloroethane (DDT) in coastal habitats by the early 1980s, followed by their diffusion into deep ocean environments. The POPs then persisted at lower concentrations. However, even recalcitrant chemicals such as DDT may be finally if slowly fading from marine food chains (*8*).

In the past decade, other contaminants have emerged as threats to marine biota. The most problematic of those are structurally similar to the chlorinated POPs but are substituted with bromine or fluorine. Access to eggs archived in specimen banks and advances in analytical chemistry have enabled retrospective analyses of such emerging pollutants (*8–12*). Similar to the legacy POPs, an exponential increase was followed by a reversal of trends following regulation (*8*, *10*).

Contaminants accumulate in seabirds mainly through their diet. Lipophilic compounds are retained in body lipids and excreted slowly; their levels are usually higher in seabirds feeding on large fish than in those feeding on plankton. Similarly, heavier nitrogen isotopes are mostly retained in body protein and, like contaminants, enriched between prey and predator. Nitrogen isotope ratios can therefore help distinguish historical changes in contamination from changes in diet (*5–10*, *14*). An increase in both nitrogen isotope ratios and contaminants over time, for instance, could indicate that the higher contamination is due to a switch in diet to more contaminated prey from a higher trophic level (*14*). Stable isotopes can also help us to understand spatial variation in contamination. For example, carbon and mercury isotope ratios demonstrated that elevated mercury levels in Alaskan seabirds originated not from upwelling or atmospheric deposition but from Yukon River sediments (*13*).

Use of miniaturized animal-borne recorders can refine the basin-scale estimates derived from isotope analysis. In one study, two species of albatross differed in accumulation of polychlorinated biphenyls (PCBs) and DDT by almost a factor of 5, despite nesting on the same atoll. Satellite tracking showed that the species with higher contamination foraged in the California Current, whereas the other species foraged off Alaska (*6*). Use of solar-powered GPS loggers with Bluetooth capability or tracked from the International Space Station could reduce battery weight and increase battery duration while eliminating the need to recapture individuals. Such technologies could track birds at meter-scale accuracy, providing more clues about the specific locations where birds derive their contaminant burdens (*15*).

Many chemical pollutants are hydrophobic and are adsorbed by plastics floating in the ocean (*4*). Seabird stomachs regularly contain hundreds of plastics; species with high plastic ingestion also have elevated contamination (*4*). If not killed directly by the garbage, seabirds with high levels of plastic may experience toxicological effects from the pollution.

The plastic not piled onto beaches or left undigested in seabird stomachs can end up circulating for decades. Likewise, toxic contaminants will tend to move toward the ocean. Long after legislation removes a chemical from commerce, that compound will continue to cycle through the ocean and be

absorbed by seabirds. It seems inevitable that there will be a need to track contaminants and assess their impact on marine wildlife long into the future.

References and Notes

1. D. W. Anderson *et al.*, *Science* **190**, 806 (1975).
2. D. Muir *et al.*, *Sci. Total Environ.* **230**, 83 (1999).
3. M. Gilbertson, J. E. Elliott, D. B. Peakall, in *The Uses of Birds*, A. W. Diamond, F. Filion, Eds. (International Centre for Birds of Prey, Newent, UK, 1987), pp. 231–248.
4. E. L. Teuten *et al.*, *Philos. Trans. R. Soc. London Ser. B* **364**, 2027 (2009).
5. A. T. Vo, M. S. Bank, J. P. Shine, S. V. Edwards, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7466 (2011).
6. M. Finkelstein *et al.*, *Ecol. Appl.* **16**, 678 (2006).
7. N. W. van den Brink, M. J. Riddle, M. van den Heuvel-Greve, J. A. van Franeker, *Mar. Pollut. Bull.* **62**, 128 (2011).
8. J. E. Elliott *et al.*, paper presented at the Annual Conference of the North Pacific Marine Science Organization (PICES), Hiroshima, Japan, 12 to 21 October 2012, abstract S6-8626; see p. 83 in www.pices.int/publications/book_of_abstracts/PICES-2012-Book-of-Abstracts.pdf.
9. B. M. Braune *et al.*, *Sci. Total Environ.* **378**, 403 (2007).
10. J. D. Crosse, R. F. Shore, K. C. Jones, M. G. Pereira, *Environ. Pollut.* **161**, 93 (2012).
11. K. E. Holmström, U. Järnberg, A. Bignert, *Environ. Sci. Technol.* **39**, 80 (2005).
12. L. B. Helgason *et al.*, *Environ. Toxicol. Chem.* **28**, 1096 (2009).
13. R. D. Day *et al.*, *Environ. Sci. Technol.* **46**, 5327 (2012).
14. C. E. Hebert, D. V. Weseloh, *Environ. Sci. Technol.* **40**, 5624 (2006).
15. E. S. Bridge *et al.*, *Bioscience* **61**, 689 (2011).

Acknowledgments: We thank C. Bishop, A. Gaston, P. Martin, and J. Provencher for useful comments, and S. Lee for assistance with drafting the figure.

10.1126/science.1235197

CANCER

Silencing a Metabolic Oncogene

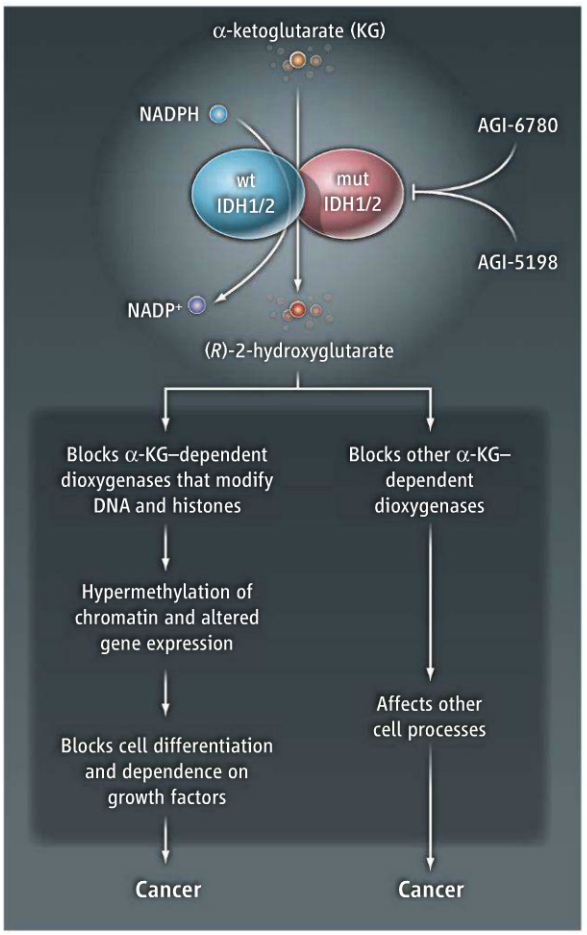
Small molecules inhibit a mutant enzyme confined to tumors, supporting therapeutic approaches that can reprogram metabolism in cancer.

Jiyeon Kim and Ralph J. DeBerardinis

Many human cancers, particularly gliomas and acute myelogenous leukemia (AML), contain mutations in the genes *IDH1* or *IDH2*, which encode two isoforms of the metabolic enzyme isocitrate dehydrogenase (*1*, *2*). These mutant enzymes produce the (*R*)-enantiomer of 2-hydroxyglutaric acid [(*R*)-2HG], a molecule that inhibits histone- and DNA-modifying enzymes, thereby altering gene expression and promoting the acquisition of malignant features (*3*–*5*). Reports by Losman *et al.* (*6*) as well as by Wang *et al.* (*7*) and Rohle *et al.* (*8*) on pages 622 and 626 of this issue, respectively, find that inhibitors of the mutant forms of IDH1/2 suppress the growth of (*R*)-2HG–producing tumor cells (*6*–*8*). The findings imply that curtailing (*R*)-2HG supply normalizes gene expression and reverses malignancy.

Metabolic reprogramming in cancer has long been considered a potential source of therapeutic targets. However, much of this reprogramming reflects the enhancement of normal metabolic activities already present in nonmalignant tissue, rather than the appearance of novel activities confined to the tumor. This makes it challenging to develop strategies that impair tumor metabolism without disturbing metabolism elsewhere. By contrast, *IDH1/2* mutations are somatically acquired and elicit an entirely new function for the enzymes (so-called "gain-of-function" or neomorphic activity) that is absent outside of the tumor (*1*–*3*). Wild-type IDH homodimers catalyze the nicotinamide adenine dinucleotide phosphate ($NADP^+$)–dependent conversion of isocitrate to α-ketoglutarate. In tumors with monoallelic mutations in *IDH1* or *IDH2*, heterodimers containing one mutant and one wild-type subunit catalyze the reduction of α-ketoglutarate to (*R*)-2HG, a reaction that depends on NADPH (the reduced form of $NADP^+$) (see the figure) (*3*, *9*). (*R*)-2HG accumulates to millimolar concentrations within the tumor (*3*). The identification of this particular metabolite as the product of mutant IDH1/2 is compelling because its (*L*)-enantiomer [(*L*)-2HG] is associated with pediatric brain tumors (*10*). These observations implicate mutant *IDH1/2*, and specifically (*R*)-2HG, as functional drivers of malignancy. More than half of "low-grade" gliomas (slow-growing but eventually lethal) and almost 10% of AML cases contain *IDH1/2* mutations, and a number of other tumors (including chondrosarcomas and cholangiosarcoma) also harbor mutations in these genes.

A key insight into the role of *IDH1/2* in cancer was that (*R*)-2HG interferes with dioxygenases that use α-ketoglutarate as a cosubstrate. These include enzymes that chemically modify histone proteins and DNA to orchestrate gene expression.



Reversing the perfect storm. Heterodimers of wild-type (wt) and mutant (mut) subunits of the metabolic enzymes IDH1/2 catalyze the production of (*R*)-2-hydroxyglutarate. Its accumulation impairs α-ketoglutarate–dependent dioxygenases, including those that modify DNA and histones (including the demethylation of 5-methylcytosine by TET-family hydroxylases and histone demethylases of the JmjC domain-containing family). This alters the epigenetic landscape, thereby blocking cell differentiation and promoting the acquisition of malignant features. Inhibitors (AGI-6780 and AGI-5198) that block (*R*)-2 hydroxyglutarate–producing IDH isoforms limit the growth of glioma- and AML-derived cancer cells.

CREDIT: C. BICKEL/SCIENCE

Children's Medical Center Research Institute, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA. E-mail: ralph.deberardinis@utsouthwestern.edu

Tumors with *IDH1/2* mutations contain a distinctive profile of DNA and histone hypermethylation and express a suite of genes associated with undifferentiated progenitors (*2*, *11*). Losman *et al*. found that exogenous expression of mutant enzymes, or treatment with (*R*)-2HG, is sufficient to reprogram epigenetics, impair differentiation, and promote malignant features in cultured erythroleukemia cells. Thus, the prevailing model is that *IDH1/2* mutations provide an oncogenic function by producing a metabolite, (*R*)-2HG, that arrests differentiation and creates a cellular context conducive to malignancy.

The suitability of these mutant enzymes as therapeutic targets hinges on whether cutting off the (*R*)-2HG supply is sufficient to induce cell differentiation and/or slow growth. Losman *et al*. demonstrate that reducing (*R*)-2HG production in erythroleukemia cells with a chemical inhibitor eliminated growth factor–independent proliferation and restored the expression of cellular differentiation markers, revealing the reversibility of the effects of mutant *IDH1* and (*R*)-2HG (*6*). Wang *et al*. developed a small-molecule allosteric inhibitor of IDH2/R140Q (IDH2 containing the common mutation $Arg^{140} \rightarrow Gln$) and demonstrated its selectivity against mutant homodimers and mutant-containing heterodimers of the enzyme. When used to treat leukemia cells, the inhibitor reduced the amount of (*R*)-2HG produced and activated the expression of genes associated with erythroid differentiation. Furthermore, in primary human AML cells, this inhibitor suppressed cell growth, reduced numbers of immature blast cells, and increased differentiation along macrophage and granulocytic lineages. None of these effects were observed in leukemic cells lacking the IDH2/R140Q mutation.

Rohle *et al*. observed that an inhibitor against IDH1/R132H (IDH1 containing the mutation $Arg^{132} \rightarrow His$), the most common *IDH* mutation in gliomas, suppressed colony formation and xenograft growth of cells from a human anaplastic oligodendroglioma, and induced the expression of genes associated with differentiation into mature glial cells. At high oral doses in mice, inhibition of IDH1/R132H reduced some histone methylation marks [whose removal is blocked by (*R*)-2HG]. Surprisingly, a lower dose of the drug impaired tumor growth but had no effect on differentiation or methylation signatures, which suggests that reversal of the epigenetic program induced by (*R*)-2HG is unnecessary to suppress glioma growth in this model. Thus, inhibiting enzymes that directly modify histones and DNA may not be equivalent to inhibiting mutant IDH1.

The appearance of *IDH1/2* mutations early in the progression of glioma and AML raised concern that (*R*)-2HG functions in tumor initiation but is dispensable once a more durable transformed state is established by the acquisition of additional mutations. It is therefore noteworthy that inhibitors to IDH1/R132H or IDH2/R140Q were effective against cells with multiple mutations. The data argue that (*R*)-2HG functions in tumor maintenance and support *IDH1/2* mutants as practical therapeutic targets.

The inhibitors used by Wang *et al*. and Rohle *et al*. produced cytostatic rather than cytotoxic effects, in line with the idea that they stimulate cell differentiation rather than cell death. If the inhibitors induce a permanent state of differentiation, perhaps no cytotoxicity is needed to achieve therapeutic efficacy. However, the survival of viable cells still containing a potentially transforming constellation of mutations makes it important to determine whether the therapeutic effects will persist over long time frames.

These studies underscore the complexity of IDH1/2 function in neoplasia. Glioma and AML progenitors are susceptible to *IDH1/2* mutations, whereas cells giving rise to many other types of cancer seem to be relatively impervious. Perhaps the basis for this selectivity relates to tissue-specific roles for the panoply of α-ketoglutarate–dependent dioxygenases in tumor suppression. Even in glioma, non-epigenetic effects of *IDH1/2* mutations appear to contribute substantially to tumor growth (*8*). It will undoubtedly be useful to understand the full scope of these effects and to maximize their suppression in cancer therapy.

References and Notes

1. H. Yan *et al*., *N. Engl. J. Med.* **360**, 765 (2009).
2. M. E. Figueroa *et al*., *Cancer Cell* **18**, 553 (2010).
3. L. Dang *et al*., *Nature* **462**, 739 (2009).
4. W. Xu *et al*., *Cancer Cell* **19**, 17 (2011).
5. C. Lu *et al*., *Nature* **483**, 474 (2012).
6. J. A. Losman *et al*., *Science* **339**, 1621 (2013).
7. F. Wang, L. Wang, A. V. Filippenko, T. Zhang, X. Zhao, *Science* **340**, 622 (2013); 10.1126/science.1234769.
8. D. Rohle *et al*., *Science* **340**, 626 (2013); 10.1126/science.1236062.
9. B. Pietrak *et al*., *Biochemistry* **50**, 4804 (2011).
10. M. Kranendijk, E. A. Struys, G. S. Salomons, M. S. Van der Knaap, C. Jakobs, *J. Inherit. Metab. Dis.* **35**, 571 (2012).
11. S. Turcan *et al*., *Nature* **483**, 479 (2012).

Acknowledgments: R.J.D. is supported by the National Cancer Institute (grant R01 CA157996), the Robert A. Welch Foundation (grant I-1733), and a Damon Runyon Cancer Research Foundation Clinical Investigator Award.

10.1126/science.1238523

CELL BIOLOGY

Unconventional Secretion, Unconventional Solutions

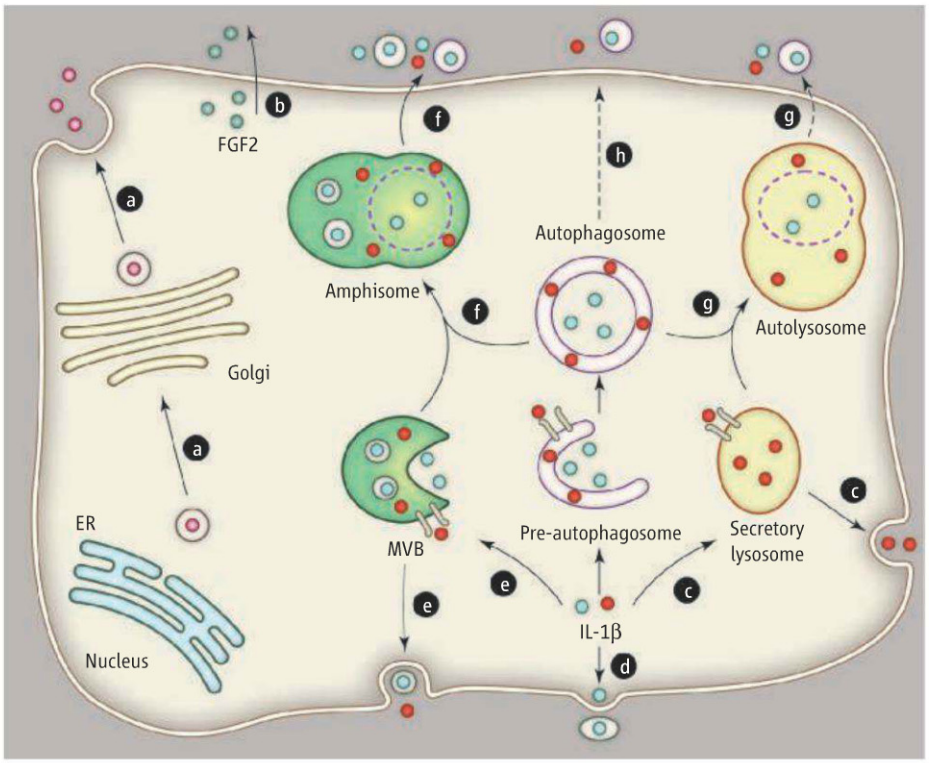
Min Zhang and Randy Schekman

Cargoes may be secreted across the plasma membrane by diverse alternative pathways.

Most eukaryotic secretory proteins are directed into the endoplasmic reticulum (ER) by an amino-terminal signal peptide (leader sequence) and progress to the cell surface through vesicular flow via the Golgi apparatus (*1*, *2*). However, many claims have been made for a number of diverse membrane and soluble proteins that lack a typical signal peptide being transported from the cytoplasm or nucleus independently of this classical ER-Golgi route, so-called unconventional secretion (*3*–*5*). Some examples, such as secretion of the yeast a-factor mating pheromone by a plasma membrane peptide transporter (*6*), are well understood. However, other examples are of proteins that are inefficiently secreted, and the concern is that in some experiments this "export" is due to cell lysis rather than actual secretion (*7*). Here we summarize the key evidence for unconventional secretion, and pose questions that remain to be addressed.

Unconventional secretion is proposed to occur by way of several different pathways (see the figure). Some leaderless cargoes may directly traverse the plasma membrane whereas others may associate with various secretory vesicles and be discharged by membrane fusion at the cell surface (*4*). The "gold standard" for demonstrating unconventional secretion directly across the plasma membrane remains the identification of a cell

Department of Molecular and Cell Biology, Howard Hughes Medical Institute, University of California, Berkeley, CA 94720–3370, USA. E-mail: schekman@berkeley.edu



Possible routes of protein secretion. **(a)** Classical protein secretion. **(b)** Direct transit across the plasma membrane independently of a cell surface transporter. (c to h) Unconventional secretion pathways by organelle carriers. **(c)** Secretory lysosome-mediated plasma membrane fusion. **(d)** Plasma membrane budding or shedding. **(e)** Exosomes generated in MVBs and discharged by fusion at the plasma membrane. **(f)** Amphisome-mediated secretion. **(g)** Direct fusion of autolysosomes with the plasma membrane. **(h)** Direct fusion of autophagosomes with the plasma membrane. Blue dots represent cargoes trapped within interior membranes. Red dots are cargoes that directly translocate across the membranes of precursor organelles.

surface transporter. ABC transporters participate in the secretion of a-factor in *Saccharomyces cerevisiae*, m-factor in *Schizosaccharomyces pombe*, and a germ cell attractant in *Drosophila melanogaster* (*5*, *6*). By contrast, fibroblast growth factor 2 (FGF2) is proposed to transit across the plasma membrane driven by phosphatidylinositol 4,5-bisphosphate–dependent protein oligomerization, tyrosine phosphorylation, and high-affinity binding to extracellular heparin sulfate proteoglycans, with no intervening membrane channel protein (*8*). Similar proposals have been made for the secretion of HIV-1 Tat, annexin A2, and FGF1 (*5*).

Other unconventional secretion pathways have been proposed to use organelle carriers. Cargoes transported by these pathways include interleukin family protein IL-1β, IL-18, yeast acyl–coenzyme A (CoA) binding protein Acb1 (AcbA in *Dictyostelium discoideum*), nuclear protein HMGB1, and engrailed homeoprotein (*4*, *9*). These less well understood processes may depend on vesicular intermediates derived from the autophagic membrane, the endosome [or multivesicular bodies (MVBs) derived from an endosome], and possibly a secretory lysosome (see the figure) (*5*). IL-1β represents an instructive example of unconventional secretion by an organelle carrier. The generation of mature IL-1β was reported to occur within lysosomes that may be induced to fuse with the plasma membrane in a Ca^{2+}- and adenosine 5′-triphosphate (ATP)–dependent event (*10*). A subsequent study suggested that ATP-dependent activation of P2X7 receptors (plasma membrane ion channels) mediates shedding of plasma membrane vesicles that are positive for extracellular IL-1β (*11*), but later results did not support this view (*12*, *13*). Yet other work supported a role for exosomes generated in MVBs as a vehicle for secretion of IL-1β (*14*). However, these studies did not localize IL-1β within MVBs or exosomes either before or after secretion. Thus, no firm conclusion of a single pathway or indeed of authentic secretion of IL-1β was established. Instead, P2X7-induced cell death and cell lysis could have explained this apparently unconventional secretion. An obvious problem is that none of the factors directly involved in IL-1β secretion has been reported.

One intriguing recent study implicated autophagy in the secretion of IL-1β and of another cytokine, IL-18, as well as the nuclear protein HMGB1 (*15*). And even before these studies, genetic evidence in yeast implicated autophagosomes and MVBs in the secretion of the acyl-CoA binding protein Acb1 (*16*, *17*). Both the yeast and the mammalian processes appear to require Golgi membrane binding proteins from the GRASP family. Further, a connection between GRASP and autophagy was made in a yeast study on the Grh protein (the yeast homolog of GRASP), which localizes to an autophagosome-like structure, compartment for unconventional protein secretion (CUPS) (*18*). The localization of a transit intermediate within autophagosomes (or in CUPS in yeast) has yet to be made, however. Thus, it remains unclear if autophagosomes or MVBs convey unconventional cargo directly to the cell surface. It is equally plausible that an interruption in organelle maturation interferes in an indirect manner with unconventional secretion.

If autophagosomes or MVBs do physically convey unconventional cargo, several possible mechanisms may be considered (see the figure). Cargo molecules could be engulfed by an autophagosome or taken up within invaginations at the surface of an endosome that give rise to intralumenal vesicles. In this case, cargoes are trapped within the interior membrane of autophagosomes or in exosomes of MVBs. A direct discharge of autophagosomes or exosomes at the cell surface would result in vesicle-enclosed cargo. Alternatively, autophagosomes or MVBs may fuse to form amphisomes or fuse with specialized secretory lysosomes where the internal membrane may be permeabilized or dissolved, releasing soluble content into the lumen. Subsequent fusion of the mixed compartment at the plasma membrane would result in the release of trapped cytosolic components to the cell exterior. However, this requires the action of a lytic agent, possibly activated by low pH, which selectively breaches the inner membrane of the autophagosome or the exosomal membrane of the MVB without rupturing the limiting membrane of the amphisome, autolysosome, or MVB. Under normal circumstances, vesicles delivered from MVBs are selectively degraded within the lysosome. However, it is not clear how a protease-sensitive molecule such as IL-1β could survive the harsh interior of a normal lysosome.

Another possibility is that molecules directly translocate across the membrane of a precursor organelle, through the bilayer as proposed for FGF2 (*8*), or by a dedicated import channel as demonstrated for yeast

a-factor (*6*) (see the figure). Such a channel could be a resident membrane protein of the pre-autophagosomal organelle (or MVB) or be delivered to that structure from some other membrane during the maturation of the autophagosome. An analogous channel has been proposed to promote the entry of chaperone-assisted autophagy cargo into the lysosome (*19*). In this model, cargoes reside in the intermembrane space between the outer and inner membrane (exosomal membranes) of the autophagosome (or MVB). Imported cargo would then be discharged in soluble form to the cell exterior concomitant with organelle fusion at the cell surface.

Two key questions must be answered to address the alternative models for unconventional secretion. One is whether the leaderless cargoes within an autophagosome or amphisome/autolysosome are conveyed directly to the cell surface. A temporal precursor-product relationship between autophagosome-bound and secreted material has not been established. It remains possible that the effect of mutations in the autophagy-related genes (Atg) or Atg protein depletion on secretion is indirect; for example, the autophagic organelle may not carry secretory cargo but instead traffic the machinery for direct translocation of unconventional cargo from the cytosol through the plasma membrane. In the case of a direct role, there is a question about the topologic location of the leaderless cargo within the autophagosome or subsequently in the amphisome/autolysosome, which must be established by visualization or biochemical fractionation. Both tests require knowledge of the origin of the autophagosome membrane, and specifically the structure responsible for unconventional secretion.

Although a unified picture of the pathways and mechanism of unconventional secretion remain elusive, the existence of this process now seems fairly secure. In the future, the application of genetic and biochemical analysis should prove crucial in identifying the essential machineries for unconventional secretion.

References

1. G. Schatz, B. Dobberstein, *Science* **271**, 1519 (1996).
2. M. C. Lee, E. A. Miller, J. Goldberg, L. Orci, R. Schekman, *Annu. Rev. Cell Dev. Biol.* **20**, 87 (2004).
3. W. Nickel, M. Seedorf, *Annu. Rev. Cell Dev. Biol.* **24**, 287 (2008).
4. W. Nickel, C. Rabouille, *Nat. Rev. Mol. Cell Biol.* **10**, 148 (2009).
5. C. Rabouille, V. Malhotra, W. Nickel, *J. Cell Sci.* **125**, 5251 (2012).
6. J. P. McGrath, A. Varshavsky, *Nature* **340**, 400 (1989).
7. W. Nickel, *Eur. J. Biochem.* **270**, 2109 (2003).
8. W. Nickel, *Traffic* **12**, 799 (2011).
9. F. Giuliani, A. Grieve, C. Rabouille, *Curr. Opin. Cell Biol.* **23**, 498 (2011).
10. C. Andrei *et al.*, *Mol. Biol. Cell* **10**, 1463 (1999).
11. A. MacKenzie *et al.*, *Immunity* **15**, 825 (2001).
12. D. Brough *et al.*, *J. Immunol.* **170**, 3029 (2003).
13. P. A. Verhoef, M. Estacion, W. Schilling, G. R. Dubyak, *J. Immunol.* **170**, 5728 (2003).
14. Y. Qu *et al.*, *J. Immunol.* **179**, 1913 (2007).
15. N. Dupont *et al.*, *EMBO J.* **30**, 4701 (2011).
16. J. M. Duran, C. Anjard, C. Stefan, W. F. Loomis, V. Malhotra, *J. Cell Biol.* **188**, 527 (2010).
17. R. Manjithaya *et al.*, *J. Cell Biol.* **188**, 537 (2010).
18. C. Bruns, J. M. McCaffery, A. J. Curwin, J. M. Duran, V. Malhotra, *J. Cell Biol.* **195**, 979 (2011).
19. S. Kaushik, A. M. Cuervo, *Trends Cell Biol.* **22**, 407 (2012).

10.1126/science.1234740

MATERIALS SCIENCE

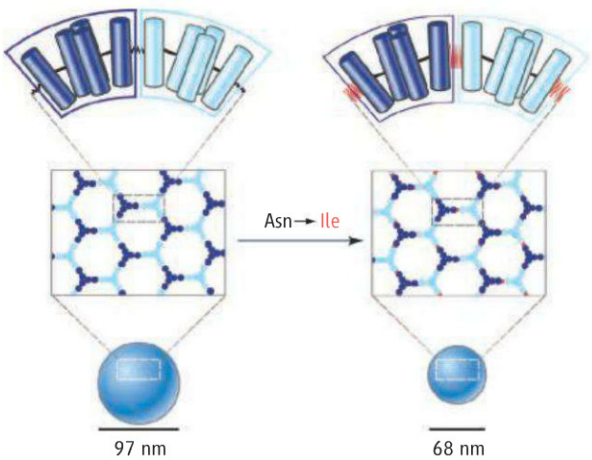
Obey the Peptide Assembly Rules

Application of simple design rules enables controlled assembly of discrete peptide nanostructures.

Maziar S. Ardejani[1,2] and Brendan P. Orner[2]

The complex properties of systems as diverse as ant colonies and stock markets emerge from the interplay of many components interacting through a small number of simple rules (*1*). Establishing the interaction rules not only helps to understand the emergence of the larger system; exploitation of the rules can provide rational control over the system itself. On page 595 of this issue, Fletcher *et al.* (*2*) describe the design of molecular components that assemble into enclosed, hollow nanostructures, the morphologies of which are defined by rules governing the interactions between the building blocks. Manipulation of these rules results in the controlled alteration of the the nanostructure's size (see the figure).

Using such rules, synthetic biology (*3*, *4*) aims to push natural biological systems in novel directions or to generate biomimetic systems with new properties. Synthetic attempts to construct biomimetic structures and systems often rely on engineering assemblies of DNA (*5*) and proteins or peptides (*6*) to adopt new structures with potentially novel functions. In living cells, proteins act as scaffolds for dynamic nanoassemblies essential for cellular structural integrity. Learning from protein noncovalent binding motifs, researchers have used short peptide structures to make self-assembled materials. The assembly properties of the peptides are governed by how their folding results in the projection of chemical functional groups into space.



Emergent morphology. In Fletcher *et al.*'s study, discrete capsular morphologies emerge through implementation of three simple rules imposed on their peptide building blocks. Subtle manipulation of the helix-helix binding through a single amino acid change (Asn → Ile) causes a drastic change in nanostructure size.

Cyclic, β sheet, and α-helical peptides are among the protein-like fragments commonly used as building blocks for assembled materials ranging from gels to mesoscopic objects (*7*). Some of the first peptide nanotubes were developed in the 1990s based on the precise assembly of cyclic peptides made up of alternating D– and L–amino acids (*8*).

Fletcher *et al.* use three design rules to control the assembly of peptide-based building blocks into discrete morphologies. The building blocks consist of three "interaction helices" radiating from a central hub. The first rule involves noncovalent binding interactions designed to link the building blocks through heterodimeric coiled coils (*9*) formed by the interaction heli-

[1]Division of Chemistry and Biological Chemistry, Nanyang Technological University, Singapore. [2]Department of Chemistry, King's College London, London, UK. E-mail: brendan.orner@kcl.ac.uk

ces. Second, the three-fold symmetry of the building blocks is established by the central hub, which is made up of a three helix bundle to which the interaction helices are projected by means of a disulfide bond. Third, the wedge shape of the building blocks in cross section results from charge repulsion at the N-termini of the central hub and the interaction helices.

Coiled-coil binding between α-helices has been used in other assembling molecular systems. However, designs for nanoscale peptide assemblies often result in discrete structure in one or two dimensions but are polydisperse overall because they aggregate continuously in a third dimension. This lack of control over uniformity is often due to design principles that fail to take into account competing, but desired processes to limit assembly. For example, one design combines heterodimeric coiled coils with overhanging "sticky ends." These peptides assemble into extended fibers of variable length (*10*). The value of coiled-coil binding to Fletcher *et al.*'s design is underlined by the fact that a subtle single amino acid change at the molecular interface results in a shift to an alternate, but still discrete, structure (see the figure). Presumably the mutation enhances binding affinity between the building blocks, resulting in an increase in the rate of closure.

Establishing symmetry in building blocks, a strategy which is incorporated into Fletcher *et al.*'s second design rule, has previously been shown to be key for the generation of other closed, hollow nanostructures, albeit small ones, using designed proteins. For example, King *et al.* induced natively symmetric multicomponent proteins to form monodisperse, hollow protein cages by using the protein structures as hubs upon which protein-protein interactions between the hubs were grafted (*11*).

The wedge shape of Fletcher *et al.*'s building block is also an essential design rule. Not unlike the size of a lipid head group influencing the convex shape of a bilayer (*12*), or the angle of the keystone affecting the diameter of an architectural arch (*13*), the skew of the wedge shape defines the curvature of the resulting enclosed nanostructure. Equally important, the wedge shape provides a means to limit the assembly as it forms. The other two design rules allow symmetric assembly but could result in the generation of extended sheets of undefined size. Curving these sheets increases the likelihood that the high-energy edges of the sheets meet, thus generating closed and highly monodisperse structures.

The three simple design rules defining the interactions in Fletcher *et al.*'s system result in the formation of a robustly controlled morphology (see the figure). Because the system is so precisely defined from the molecular to the nanoscale level, it provides easily testable mechanistic hypotheses and predictions for the controlled manipulation of the rules to generate new morphologies (*14*). It could provide a robust model system to understand how nanoscale properties emerge from molecular components. Moreover, the structures could be used for protein delivery or as biomimetic mini-organelles to sequester and transport enzymatic activity.

References

1. S. Johnson, *Emergence: The Connected Lives of Ants, Brains, Cities, and Software* (Scribner, New York, 2001).
2. J. M. Fletcher *et al.*, *Science* **340**, 595 (2013); 10.1126/science.1233936.
3. E. H. C. Bromley *et al.*, *ACS Chem. Biol.* **3**, 38 (2008).
4. W. C. Ruder, T. Lu, J. J. Collins, *Science* **333**, 1248 (2011).
5. E. Winfree, F. Liu, L. A. Wenzler, N. C. Seeman, *Nature* **394**, 539 (1998).
6. C. J. Reedy, B. R. Gibney, *Chem. Rev.* **104**, 617 (2004).
7. J. Kim *et al.*, *J. Am. Chem. Soc.* **134**, 20573 (2012).
8. M. R. Ghadiri, J. R. Granja, R. A. Milligan, D. E. McRee, N. Khazanovich, *Nature* **366**, 324 (1993).
9. E. K. O'Shea, R. Rutkowski, P. S. Kim, *Science* **243**, 538 (1989).
10. M. J. Pandya *et al.*, *Biochemistry* **39**, 8728 (2000).
11. N. P. King *et al.*, *Science* **336**, 1171 (2012).
12. J. N. Israelachvili, D. J. Mitchell, B. W. Ninham, *Faraday Transac.* **2**, 1525 (1976).
13. M. S. Ardejani, X. L. Chok, C. J. Foo, B. P. Orner, *Chem. Commun.* **49**, 3528 (2013).
14. C. Tarabout *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7679 (2011).

10.1126/science.1237708

BIOCHEMISTRY

As Good as Chocolate

Structural details of how ligands bind to serotonin receptors should guide the development of pharmaceuticals with fewer side effects.

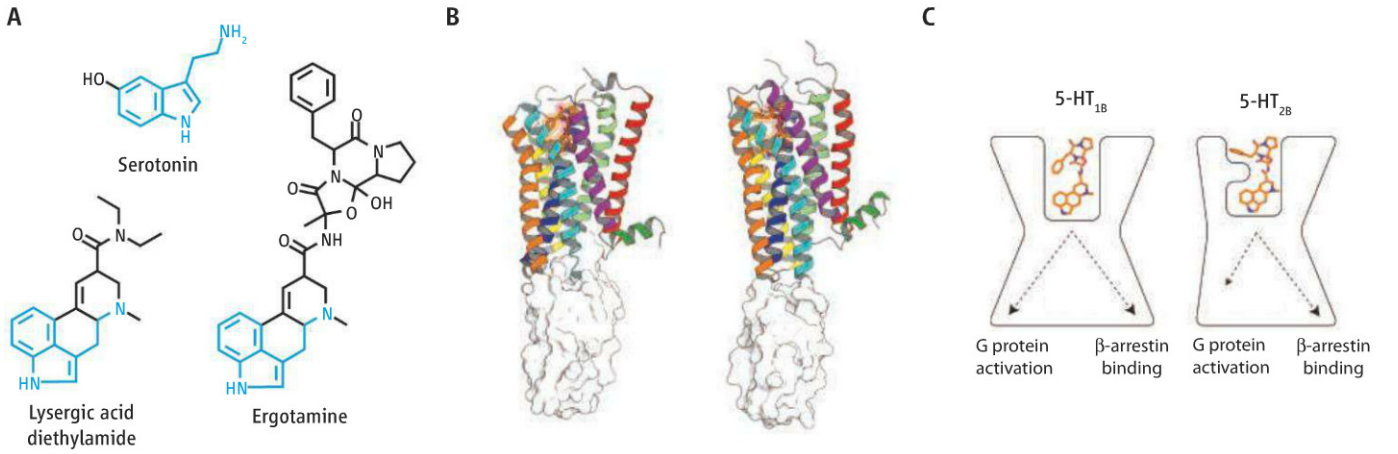
Krzysztof Palczewski and Philip D. Kiser

No one could have imagined how important the 1948 discovery of the vasoconstrictor serotonin (5-hydroxytryptamine or 5-HT) would be to the field of human physiology (*1*). Elucidation of the 5-HT structure (*2*) and synthesis of the molecule with the expected biological activity (*3*) soon followed. This monoamine is a ligand for 15 receptors, and drugs that target 5-HT receptors are widely used to treat conditions including migraine headache, depression, anxiety, nausea, vomiting, and irritable bowel syndrome, reflecting the wide diversity of physiological and pathophysiological processes in which 5-HT is involved (*4*). On page 615 and 610 in this issue, Wacker *et al.* (*5*) and Wang *et al.* (*6*), respectively, report the crystal structure of human 5-HT_{2B} bound to the antimigraine agent ergotamine and compare it with the 5-HT_{1B}-ergotamine structure. Together with biochemical and computational data, these structures reveal molecular mechanisms responsible for divergent signaling patterns of ergotamine, serotonin, and the psychedelic drug lysergic acid diethylamide (LSD).

The structures were obtained by fusing either receptor to a thermally stabilized bacterial protein [apocytochrome b562RIL (BRIL)]. This approach stabilizes the receptor to promote crystallization but does not alter ligand-binding properties. The structural information, together with computational ligand-docking experiments, reveal similar binding modes for ergotamine, 5-HT, and LSD to the ligand-binding pocket formed by residues conserved in the 5-HT receptor family, thereby clarifying the family-wide agonist activity of 5-HT. However, there are some key differences between the two receptors (see the figure). In both structures, an accessory binding pocket adjacent to the binding site for the natural ligand (5-HT) can accommodate chemical groups located distal to the core indoleamine moiety in a differential manner, which possibly could control signaling. The 5-HT_{1B} receptor displays a 3 Å outward shift at the extracellular end of helix V relative to the 5-HT_{2B} receptor, resulting in a more open, extended pocket that explains receptor subtype selectivity for ligands.

5-HT receptor subtypes are classified according to their ligand-binding preferences, sequence homology, and signaling mechanisms. With the exception of the type 3 recep-

Department of Pharmacology, School of Medicine, Case Western Reserve University, 2109 Adelbert Road, Cleveland, OH 44106–4965, USA. E-mail: kxp65@case.edu; pdk7@case.edu



Uncovering bias in serotonin receptor signaling. (**A**) Natural (serotonin/5-HT) and synthetic (ergotamine and LSD) serotonin receptor ligands have a common indole amine structure (blue) at their core. (**B**) Crystallographic structures of the human 5-HT_{1B} (left) and 5-HT_{2B} (right) receptors in a complex with ergotamine (orange). The BRIL fusion protein used to facilitate crystallization is shown as a white surface. (**C**) Ergotamine stabilizes distinct conformations in the two 5-HT receptors, providing a structural explanation for the biochemically observed phenomenon of biased signaling.

tor, which is a ligand-gated cation channel, all 5-HT receptors are seven-transmembrane receptors belonging to the rhodopsin-like class A group that transmit signals to heterotrimeric GTP-binding proteins (G proteins) and other effectors like β-arrestins (*7*). Many 5-HT receptors display a close evolutionary relationship with receptors for other biogenic amines (such as the neurotransmitters dopamine and norepinephrine). This, along with structural similarities among the biogenic amines themselves, explains why drugs that target specific 5-HT receptors are especially prone to produce untoward side effects through "off target" receptors. One example is the severe vasoconstriction that can result from supratherapeutic doses of ergot alkaloids, which are thought to exert their beneficial antimigraine effects through 5-HT_{1B}, and 5-HT_{1D} receptors but also can act on α-adrenergic receptors with dangerous consequences (*8*). Another prominent example is the fibrotic cardiac valvulopathy (thickening of heart valves) associated with a metabolite of fenfluramine that can act on the 5-HT_{2B} receptor. Fenfluramine is a component of the notorious and now defunct anorexogenic drug combination Fen-Phen (*9*), whose recall was the largest in U.S. history (*10*).

Rational drug design, in which the known structure of an endogenous agonist of interest is used as a basis to generate derivatives with potential selectivity for a subset of receptors, has played an important role in developing more selective 5-HT receptor–targeting drugs with improved side-effect profiles (*11*). A prime example is the 5-HT_1–selective "triptans," which are the drug class of choice for treating migraine headache. The 5-HT receptor structures will guide the tailoring of candidate drug molecules to bind selectively to particular receptor subtypes. Differences in the ligand-binding pockets of the 5-HT_{1B} and 5-HT_{2B} receptors could, for example, be exploited to eliminate agonist effects at the 5-HT_{2B} site that are associated with cardiotoxicity.

At first glance, structures of these receptors appear highly similar to that of rhodopsin, the first G protein–coupled receptor (GPCR) whose structure was solved by x-ray crystallography (*12*). Indeed, pairwise root mean square deviations (RMSDs) between the 5-HT receptors and rhodopsin are ~2.3 to 2.7 Å among the core of ~260 Cα positions (or ~80% of the receptor sequence), demonstrating a structural triumph whereby the same overall GPCR topology is maintained despite markedly different amino acid sequences (*13*). However, the differential ligand- and effector-binding specificity of the structures provides an important pharmacological story. Wacker *et al.* and Wang *et al.* could determine the 5-HT_{1B} and 5-HT_{2B} receptors, respectively, in the presence of the same ligand. By comparing the two structures, they noted differences in the intracellular region that interacts with G proteins or arrestins. Both receptors show biased signaling, in which agonists preferentially activate one pathway over the other. The comparison revealed that ergotamine biases the signaling of 5-HT_{2B} through β-arrestin by inducing conformational changes within the cytoplasmic portion of the receptor that are distinct from changes observed in the 5-HT_{1B} structure that enable coupling to G proteins. Such structural insights can open a whole new avenue of investigation of ligand-induced differential signaling.

The serotonin receptor family is like the mythical Roman god Janus, often depicted with two faces pointed in opposite directions. These receptors can be dangerous, like 5-HT_{2B}, which is often referred to as a death receptor because of its cardiotoxic effects. On the other hand, chocolate, which contains high concentrations of the serotonin precursor tryptophan and serotonin-like monoamines, elicits cravings through these same receptors and brings much pleasure to our lives (*14*). Structural information about GPCRs continues to have important implications for developing new pharmaceuticals (*15*). We can look forward to the design of noncardiotoxic 5-HT receptor ligands that will make the outcome of stimulating these receptors as good as chocolate.

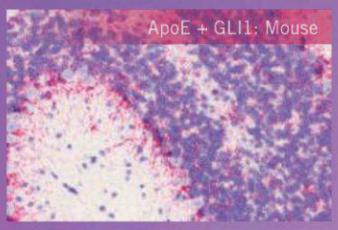
References and Notes

1. M. M. Rapport, A. A. Green, I. H. Page, *Science* **108**, 329 (1948).
2. V. Erspamer, B. Asero, *Nature* **169**, 800 (1952).
3. G. Reid, M. Rand, *Nature* **169**, 801 (1952).
4. M. Berger *et al.*, *Annu. Rev. Med.* **60**, 355 (2009).
5. D. Wacker *et al.*, *Science* **340**, 615 (2013); 10.1126/science.1232808.
6. C. Wang *et al.*, *Science* **340**, 610 (2013); 10.1126/science.1232807.
7. B. L. Roth, *Neuropharmacology* **61**, 348 (2011).
8. W. J. Meyler, *Cephalalgia* **16**, 5 (1996).
9. D. R. Boughner, *Can. Med. Assoc. J.* **157**, 705 (1997).
10. G. Cowley, K. Springen, *Newsweek* **130**, 46 (1997).
11. P. P. Humphrey, *Headache* **47** (suppl. 1), S10 (2007).
12. K. Palczewski *et al.*, *Science* **289**, 739 (2000).
13. T. Mirzadegan *et al.*, *Biochemistry* **42**, 2759 (2003).
14. K. Bruinsma, D. L. Taren, *J. Am. Diet. Assoc.* **99**, 1249 (1999).
15. J. A. Salon, D. T. Lodowski, K. Palczewski, *Pharmacol. Rev.* **63**, 901 (2011).

Acknowledgments: This work was supported by NIH grant EY008061.

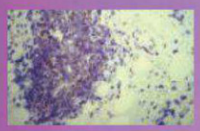
10.1126/science.1238521

CREDIT: C. BICKEL/SCIENCE

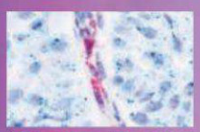


No Antibody.
No Problem.

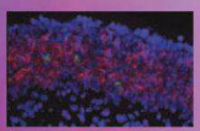
Get immediate results with RNAscope® ISH Assays for any gene in any species.



RNAscope Single-Plex ISH
FGFR1 mRNA
FFPE breast cancer tissue



RNAscope 2-Plex ISH
PECAM1 and EGFR mRNA
FFPE breast cancer tissue



RNAscope Multiplex ISH
OMP and OLFR73 mRNA
Fresh-frozen olfactory bulb

Exquisite sensitivity
Custom probes in < 2 weeks
Over 27,000 predesigned probes
Complete the assay and see your data in one day
Easy and guaranteed to work

www.acdbio.com/no-antibody

Be Amazed.

ACD

Join the Conversation!

Twitter is a great way to connect with AAAS members and staff about the issues that matter to you most. Be a part of the discussion while staying up-to-date on the latest news and information about your personal member benefits.

Follow us @AAASmember

and join the conversation

with #AAAS

AAAS MEMBERCENTRAL

MemberCentral.aaas.org

INTRODUCTION

Alien Worlds Galore

WE KNEW THEY HAD TO BE THERE. ONCE PEOPLE REALIZED THAT THE SUN IS A STAR and that the planets travel around it, it was only natural to suspect that other stars had planets of their own. Confirmation, however, didn't come until the early 1990s, when astronomers found compelling signs of smaller bodies orbiting first pulsars and then normal main-sequence stars.

Today, it is estimated that our galaxy contains at least as many planets as stars. Almost 900 have been detected, and a few thousand others are under investigation. Observations have shown that planets are not only common but also extremely diverse, with our solar system being just one example of this diversity and perhaps even a rather unusual one. As described by Howard (p. 572), the most commonly observed class of planetary system so far consists of one or more planets one to three times the size of Earth, orbiting much closer to their star than Earth does to the Sun.

Over the next decade, astronomers plan to find and characterize habitable planets. William Borucki *et al.* (p. 587) have taken a step toward this goal by detecting two small planets (1.4 and 1.6 times the size of Earth) in the habitable zone of a star smaller and cooler than the Sun. In this Report, the habitable zone is defined as "the annulus around a star where a rocky planet with a CO_2-H_2O-N_2 atmosphere and sufficiently large water content (such as on Earth) can host liquid water on its solid surface." Seager (p. 577) argues that the habitable zone concept needs to be expanded to include other possibilities such as hydrogen-rich atmospheres and reduced water inventories. Given the expected diversity of exoplanet atmospheres and interiors, she suggests that habitability must be considered on a case-by-case basis rather than as a one-size-fits-all definition if we are to increase our future chances of detecting chemical signatures of life. Finding these signatures will be difficult, however, and once they are detected, we may not be 100% certain that they were produced by life.

Meanwhile, in the field of planetary discovery itself, the signs of life could not be stronger. Every month, the orbiting Kepler telescope alone sights hundreds of new potential exoplanets in a patch of sky near the constellation Cygnus. Several more ground- and space-based missions are either in progress, in the works, or under consideration. To help nonexperts navigate this brave new world of world-finding, News writers Yudhijit Bhattacharjee and Daniel Clery (p. 566) survey some of the main techniques and research efforts in line for the next round of planetary discoveries, and Lizzie Wade (p. 570) presents a not-too-technical guide to exo-nomenclature. In the News Focus section (p. 542), Bhattacharjee profiles Kepler's creator, Borucki.

Now for the next step: finding things we *don't* know are there—discoveries that will quickly render this special section out of date.

– MARIA CRUZ AND ROBERT COONTZ

Exoplanets

CONTENTS

News

566 Gallery of Planet Hunters
570 … And a Glossary of Their Quarry

Reviews

572 Observed Properties of Extrasolar Planets
A. W. Howard
577 Exoplanetary Habitability
S. Seager

See also News story p. 542, Report p. 587,
Science *Podcast, Planet Hunters slideshow,*
and more at www.sciencemag.org/special/exoplanets

Science

CREDIT: ESO/T. PREIBISCH

A Gallery of Planet Hunters

–YUDHIJIT BHATTACHARJEE AND DANIEL CLERY

In 1992, astronomers announced that a whirling neutron star called a millisecond pulsar harbored the first known planets outside our solar system: extrasolar planets, or exoplanets. Two decades on, exoplanets are practically an astronomical commodity. Nearly 900 of them have been confirmed, and hundreds of fresh candidates are turning up every month. Some resemble the familiar orbs circling our sun, but others are truly alien: hot Jupiters, mini-Neptunes, super-Earths.

To help readers navigate the bewildering variety of new worlds and their discoverers, *Science* presents this guide for the perplexed: an overview of planet-hunting techniques and representative efforts, followed by an informal glossary of popular exo-objects.

Radial Velocity

Radial velocity measurement is the mother of exoplanet search techniques. Michel Mayor and Didier Queloz of the Observatory of Geneva in Switzerland used it in 1995 to make the first confirmed detection of an exoplanet around a sunlike star, and it's still going strong.

As a planet orbits a star, its gravity causes the star to move around their common center of gravity, usually inside the star. If we view such a system edge-on, the star would appear to be wobbling from side to side. Current telescopes can't detect such tiny movements, but they can spot a wobble back and forth along the line of sight because such changes in radial velocity raise and lower the frequency of the star's light.

To spot exoplanets this way, astronomers don't need a huge telescope or one in space—just an extremely sensitive spectrometer and lots of time. It can require between 500 and 1000 separate observations of a star to detect the signature of a planet circling around it. Today's best instruments can detect radial velocities down to about 1 meter/second. Jupiter makes our sun move at 12 m/s, but small planets cause much smaller movements—Earth shifts the sun by less than 0.1 m/s.

HARPS and HIRES

Ongoing

Many groups around the world have hitched spectrometers to their telescopes and started looking for wobbling stars, but two groups have dominated the discovery of new exoplanets with several hundred detections each. The first is the Geneva Observatory team with an instrument called High Accuracy Radial velocity Planet Searcher (HARPS) fitted to the European Southern Observatory's (ESO's) 3.6-meter telescope at La Silla in Chile. The second team, led by Geoff Marcy of the University of California, Berkeley, uses the High Resolution Echelle Spectrometer (HIRES) on the Keck I telescope at Mauna Kea, Hawaii.

Since the Kepler space mission began detecting new candidate exoplanets by the thousands using the transit technique (see p. 567), radial velocity teams have changed focus from discovering new planets to confirming Kepler detections and measuring their mass. Kepler observations can measure a planet's size but not its mass, which is key to separating gas giants from rocky planets. Radial velocity fills that gap, and so the teams are trying to shift up a gear.

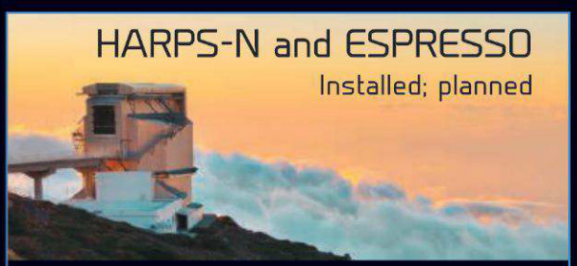
Automated Planet Finder

Imminent

Because radial velocity measurement is so slow, both teams are aiming to develop new instruments and telescopes dedicated to planet searching. The HIRES team has built the Automated Planet Finder, a 2.4-meter telescope and spectrometer on Mount Hamilton near San Jose, California. Its spectrometer is about as sensitive as that of HIRES, and starting this month it will spend all night, every night looking for planets—taking 25 spectra per night.

HARPS-N and ESPRESSO

Installed; planned



Meanwhile, the HARPS team and collaborators have made another version of their spectrometer for the northern hemisphere. HARPS-N is fixed to Italy's 3.6-meter Galileo National Telescope in the Canary Islands, which can observe the same patch of sky that Kepler does. The researchers are also designing a next-generation spectrometer called ESPRESSO, which will operate from 2016 at ESO's Very Large Telescope at Cerro Paranal in Chile. VLT is an array of four 8.2-meter scopes; ESPRESSO will be portable enough to attach to whichever has available observing time.

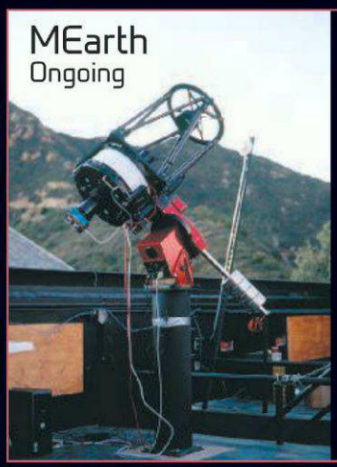
CREDITS (CLOCKWISE FROM TOP): (BANNER) ESO/T. PREIBISCH; © LAURIE HATCH/WWW.LAURIEHATCH.COM; TNG/A. HARUTYUNYAN; ESO/F. KAMPHUES

Transits

The idea is simple enough: monitoring a star for the dips in its brightness caused by a planet passing in front of it. In practice, however, such transit-based detection is like trying to tell when a seagull flies across the beam from a lighthouse.

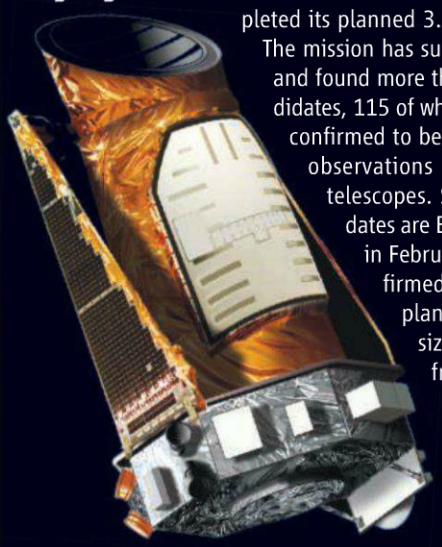
That's why it took astronomers years to convert the idea into a workable technique capable of finding not just gigantic planets, which block more than 1% of their star's light, but also planets the size of Earth. Their efforts have paid rich dividends, enabling the discovery of more than 231 exoplanets to date.

MEarth
Ongoing



On a mountaintop in Arizona, eight remote-controlled robotic telescopes search the skies every night for transits around M dwarfs. The most common stars in our galaxy, M dwarfs are much smaller than the sun, so the dip in their brightness due to a transiting planet should be larger than the dimming of a sun-sized star. In 2009, MEarth discovered its first super-Earth, a planet orbiting a red dwarf 40 light-years from us.

Kepler
Ongoing



Launched in 2009, NASA's Kepler spacecraft has entered an extended phase, having completed its planned 3.5 years of observing. The mission has surveyed 150,000 stars and found more than 2740 planet candidates, 115 of which astronomers have confirmed to be planets by follow-up observations using ground-based telescopes. Several of the candidates are Earth-sized or smaller; in February, astronomers confirmed one of them to be a planet approximately the size of our moon. Data from Kepler are helping researchers get a better handle on how abundant different kinds of planets are.

TESS
Planned

The Transiting Exoplanet Survey Satellite (TESS)—recently chosen by NASA for possible launch in 2017—will aim to survey the brightest stars in the sun's neighborhood using an array of wide-field cameras in search of exoplanets ranging from gas giants to extraterrestrial Earths in habitable zones. Some of those planets, researchers hope, will be candidates for follow-up studies of their atmospheres by the James Webb Space Telescope, scheduled for launch in 2018.

CHEOPS
Under Construction

If all goes as planned, by 2017 European astronomers will launch a spacecraft to look at transits of nearby bright stars already known to harbor planets. Characterizing Exoplanets Satellite (CHEOPS)—chosen for development by the European Space Agency in 2012—is designed to reveal in detail the nature of those planets. High-precision measurements by the satellite should help astronomers nail down planet sizes. Combined with radial velocity observations from the ground, which provide the mass of these planets, the observations will help figure out how dense they are.

CREDITS (CLOCKWISE FROM TOP RIGHT): NASA/KEPLER MISSION; SWISS SPACE CENTER/CHEOPS TEAM; CHESTER BEALS/MASSACHUSETTS INSTITUTE OF TECHNOLOGY; JONATHAN IRWIN

Microlensing

If you point a telescope at a star in the center of the galaxy and stare long enough, sooner or later a second star will pass in front of it. (Yes, it could take millions of years.) Depending on its mass and position, the foreground star can exert enough gravity to bend the light from the background star, acting as a lens. Such "microlensing" makes the background star appear brighter than usual.

A planet orbiting the foreground star can make brightening or dimming contributions of its own. Such deviations from what would otherwise be a smooth microlensing event can be picked up by planet hunters to reveal the existence of the planet. With precise observations of this signal, astronomers can determine the mass of the planet and its distance from the parent star.

Using microlensing, astronomers can detect planets much less massive than Earth, which other methods would miss. The drawback is that it's a one-shot deal: Microlensing events, lasting days to weeks, are so rare that researchers are unlikely to get a second chance to study a particular foreground star. That's why ongoing microlensing-based searches involve networks of robotic telescopes that can watch big swaths of the sky to spot as many events as possible. The technique has yielded the discovery of 17 exoplanets to date.

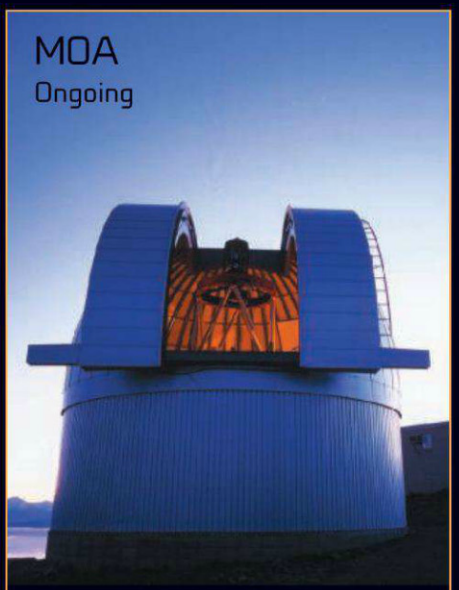
The Optical Gravitation Lensing Experiment

Ongoing

In the mid-1990s, a group of Polish astronomers led the construction of a 1.3-meter telescope at the Las Campanas Observatory in Chile. The Optical Gravitational Lensing Experiment's main goal has been to search for dark matter using microlensing, but since the early 2000s, the telescope has also proved its utility as a planet-hunting tool. Its capabilities have greatly improved since first light. In 2009, the telescope was mounted with a new, 32-chip mosaic camera, a far cry from the single-chip instrument used in its early years.

MOA

Ongoing



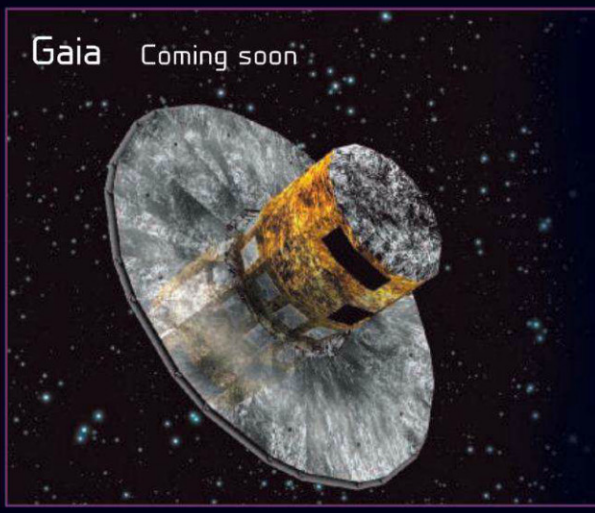
A 1.8-meter telescope at Mount John University Observatory in New Zealand has been surveying the sky for microlensing events since 2006, driven by a collaboration among astronomers in Japan, New Zealand, and the United States. Microlensing Observations in Astrophysics (MOA) has the biggest field of view of any telescope involved in microlensing surveys: an area of 2.2 square degrees of sky. Observations using MOA have led to the discovery of rogue planets: planetary bodies that appear to be gravitationally unbound from any stars.

Astrometry

Like the radial velocity technique, astrometry seeks to detect the slight wobble in a star's position caused by the gravity of an orbiting planet. Astrometry does it by plotting the star's position in the sky with extreme accuracy, through many observations repeated over time. Astronomers have been claiming the detection of "unseen companions" to stars using astrometry since 1855, but none have stood up to scrutiny. To date, there have been no confirmed exoplanet detections using the technique.

Gaia

Coming soon



Astronomers hope that things will change later this year when the European Space Agency launches Gaia, an orbiting observatory that will accurately map the positions and motions of a billion stars in our galaxy, 1% of its population. Gaia's goals include detecting tens of thousands of exoplanetary systems. Researchers hope that Gaia will tell them more about the distribution of exoplanets around the galaxy: Are there more near the center or in the spiral arms? Are planets more common in areas rich in heavy elements?

CREDITS (TOP TO BOTTOM): MOA; K. ULACZYK; ESA

Direct Imaging

The Holy Grail of exoplanet searching is direct imaging. If you can gather light from the planet itself, you can get a spectrum and find out the planet's temperature and what it's made of, compare its composition with those of its star and sister planets, and look for signs of life.

The problem is that a planet is very close to a star that is a million or more times brighter—a firefly in the glare of a searchlight. Spotting one is not impossible. A handful of planets have been observed by both ground-based telescopes and Hubble. These, however, have tended to be "low-hanging fruit": large planets with wide orbits around dim stars.

Direct imaging is all about contrast: picking out the faintest thing that can be discerned next to the star. Current systems can detect planets about a millionth as bright as their star. To do that, a telescope needs two things: a coronagraph and wavefront control. A coronagraph is simply an opaque patch that blocks the light of the star so that the dim planet can be seen. Starlight diffracted around its edge can still swamp the planet's light, so sophisticated optics are employed to ensure that this edge light is directed out of the way.

Wavefront control employs a range of techniques to improve the dynamic range of the image by correcting for optical faults in the telescope. In ground-based instruments, it is essential to include extreme adaptive optics: computer-controlled deformable mirrors that move in real time to correct for atmospheric distortion.

WFIRST

Under Consideration

Atmospheric disturbance limits what ground-based telescopes can do, so to observe Earth-like planets will require instruments in space. Both NASA and the European Space Agency have designed dedicated planet-imaging spacecraft—Terrestrial Planet Finder and Darwin—with mirrors several times the size of the Hubble Space Telescope's. But both projects have been shelved as too costly.

Hoping to get something more modest into orbit, exoplanet researchers have suggested adding a coronagraph to the proposed Wide-Field Infrared Survey Telescope (WFIRST), which was designed to look for exoplanets by gravitational microlensing (see p. 568) and to probe dark energy. WFIRST was supposed to have a 1.5-meter mirror, but that plan is being reconsidered since the U.S. National Reconnaissance Office gave NASA two 2.4-meter telescopes built for spy satellites (*Science*, 15 February, p. 748). NASA asked for proposals for using the two scopes; ideas include turning one into WFIRST and building a dedicated exoplanet imager. NASA will continue to study the proposals this year.

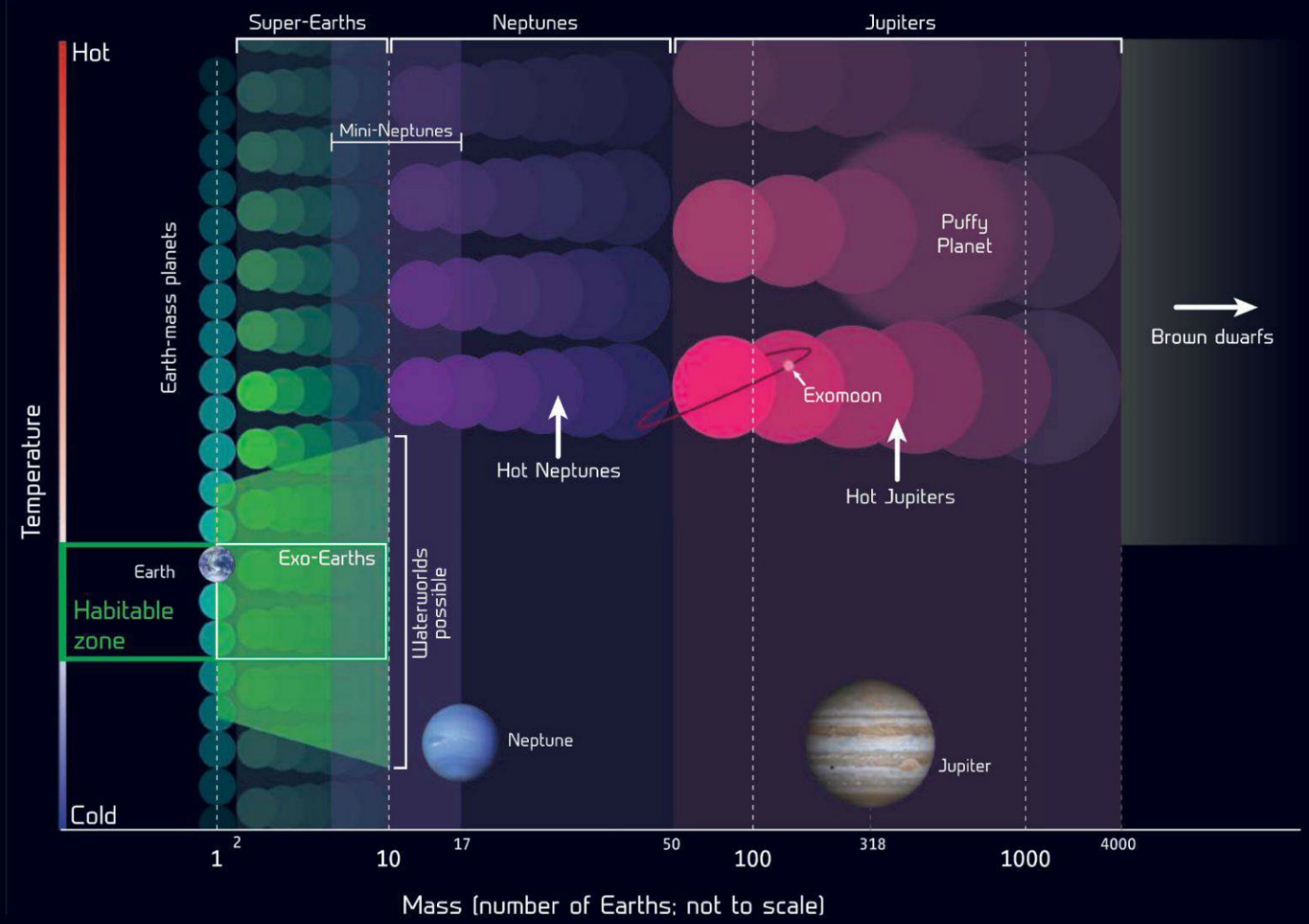
Gemini and SPHERE Coming Soon

Several projects around the world are attempting to image exoplanets. Later this year, two purpose-built instruments will begin to bring direct imaging into the mainstream by improving contrast to about 1 in 10^7. The Gemini Planet Imager is a second-generation coronagraph with adaptive optics that will be attached to the 8-meter Gemini South telescope at Cerro Pachón in Chile. SPHERE is a system being installed at the European Southern Observatory's Very Large Telescope (pictured) at Cerro Paranal in Chile that incorporates adaptive optics and three different coronagraphs. Both instruments will seek to image giant planets around nearby stars. They'll target young star systems because recently formed planets will still be hot and so will emit their own infrared light. Observed in the infrared, planets appear brighter and stars dimmer than at visible wavelengths.

CREDITS (TOP TO BOTTOM): WFIRST TEAM/NASA GODDARD SPACE FLIGHT CENTER; ESO/B. TAFRESHI/TWANIGHT.ORG

... And a Glossary of Their Quarry

–LIZZIE WADE



Hot Jupiter

Gas giants more than 50 times the mass of Earth orbiting too close to their stars to be habitable. Make up 42% of confirmed planets. Example: **51 Pegasi b**.

Puffy Planet

Similar to a hot Jupiter in mass and proximity to a star, but much less dense and thus larger. Synonym: Inflated hot Jupiter. Example: **HAT-P-1**.

Hot Neptune

Planets between 10 and 50 Earth masses orbiting close to their stars. Example: **Gliese 436 b**.

Mini-Neptune

A planet that is less massive than Neptune but shares its characteristic thick atmosphere of hydrogen and helium. Current detection techniques can have trouble telling them from super-Earths. Example: **Kepler-11 b-f**.

Super-Earth

A planet with about two to 10 times the mass of Earth. Can be hard to distinguish from a mini-Neptune. Example: **Kepler-62 e and f**.

Earth-mass planet

A planet about the mass of Earth, but not necessarily within its star's habitable zone. Example: Planet orbiting **Alpha Centauri Bb**.

Exo-Earth

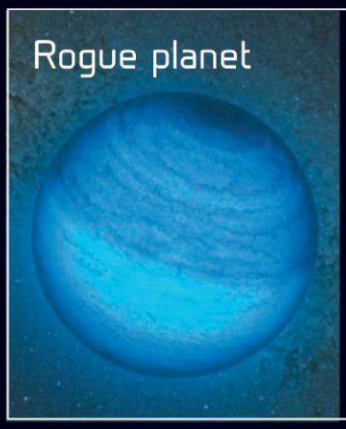
A rocky planet with one to 10 times the mass of Earth, orbiting in the habitable zone of its star. Possible example: expected any time. Synonyms: Twin Earth, Goldilocks planet, Earth 2.0, Earth analog, Earth-like planet.

CREDITS: (GRAPHIC PHOTOS LEFT TO RIGHT) NASA/APOLLO 17 CREW; NASA/JPL; NASA/JPL/UNIVERSITY OF ARIZONA; (BOTTOM) NASA/TIM PYLE

Pulsar planet

An exoplanet orbiting a pulsar, a spinning neutron star left behind after a supernova; probably debris from the explosion, trapped by the neutron star's strong gravity. Only a handful are known. Not habitable. Examples: **PSR 1257+12 b, c, and d**, the first exoplanets ever discovered back in 1992.

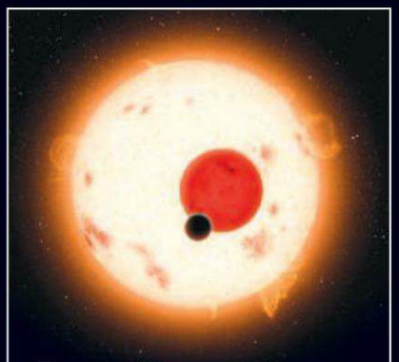
Rogue planet



A planet that does not orbit a star. Astronomers don't know whether these wandering planets were ejected from star systems or formed by themselves in interstellar space. Rogue planets can be up to 13 times the mass of Jupiter; more-massive bodies are classified as brown dwarfs. Without stars to keep them warm, they are always frigid. Synonyms: orphan planet, homeless planet, nomad planet, free-floating planet, sub-brown dwarf. Example: **CFBDSIR 2149-0403**.

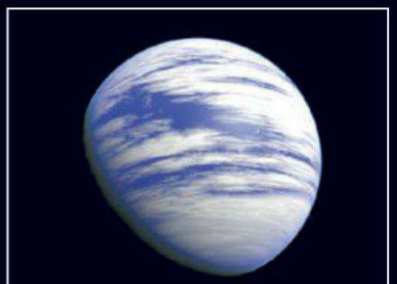
Exomoon

A moon orbiting an exoplanet. Some exomoons of gas giants may have molten interiors thanks to tidal heating, which could keep them warm even outside their stars' habitable zones and make them easy to spot from afar. But so far, no exomoons have been observed. Possible example: **Fomalhaut b**.



Circumbinary planet

A planet that orbits a binary star system and thus has two suns instead of one. Examples: **Kepler-16 b, Tatooine**.



Waterworld

A super-Earth covered in water—in the form of ice, oceans, or a water-vapor atmosphere, depending on the planet's proximity to the star. Example: **Gliese 1214b**.

Eccentric planet

With orbits the shape of stretched-out ellipses, these planets can swoop through drastically different temperature zones in the course of a year. Example: **HD 80606 b**.

Core planet

The rocky core left behind after the gas of a hot Jupiter evaporates as a result of orbiting so close to its star. Synonym: chthonian planet, evaporated remnant core. Example: **COROT-7b** (suspected); **HD 209458 b** is in the process of evaporating.

Brown dwarf

Sometimes called a failed star, a brown dwarf forms when a cloud of gas collapses but is not massive enough to ignite the fusion reactions that fuel fully formed stars. Brown dwarfs can be 13 to 75 times the mass of Jupiter. Less massive bodies are rogue planets; more massive ones can sustain fusion and are thus stars. Some brown dwarfs have their own planets. Example: **2M1207**.

CREDITS (CLOCKWISE FROM TOP LEFT): NASA/JPL-CALTECH/R. HURT/SSC; ESO/L. CALÇADA/P. DELORME/R. SAITO/VVV CONSORTIUM; NASA/JPL-CALTECH; NASA/ESA/D. AGUILAR/HARVARD-SMITHSONIAN CENTER FOR ASTROPHYSICS; ESO/L. CALÇADA; NASA/JPL; ESA/NASA/L. CALÇADA/ESO FOR STSCI.

REVIEW

Observed Properties of Extrasolar Planets

Andrew W. Howard

Observational surveys for extrasolar planets probe the diverse outcomes of planet formation and evolution. These surveys measure the frequency of planets with different masses, sizes, orbital characteristics, and host star properties. Small planets between the sizes of Earth and Neptune substantially outnumber Jupiter-sized planets. The survey measurements support the core accretion model, in which planets form by the accumulation of solids and then gas in protoplanetary disks. The diversity of exoplanetary characteristics demonstrates that most of the gross features of the solar system are one outcome in a continuum of possibilities. The most common class of planetary system detectable today consists of one or more planets approximately one to three times Earth's size orbiting within a fraction of the Earth-Sun distance.

Extrasolar planets can be detected and characterized by means of high-resolution spectroscopy or precision photometry of the stars that they orbit. Although some planetary systems have familiar properties, many have characteristics not seen in the solar system—small star-planet separations that result in planets heated to >1000 K, highly eccentric and inclined orbits, planets orbiting binary stars, and planet masses and sizes not represented locally. This diversity of planetary systems echoes the Copernican principle: Earth is not the center of the universe, and the solar system does not provide a universal template for planetary system architectures.

Measurements of the properties of a large number of planetary systems can probe the mechanisms of planet formation and place our solar system in context. These surveys can answer questions such as, What are common planet sizes and architectures? and, How did such planetary systems form? Measurements of extrasolar planets are mostly limited to gross physical properties, including mass, size, orbital characteristics, and in some cases composition. Detailed measurements made in the solar system—such as spatially resolved imaging, in situ observations, and sample return—are infeasible for extrasolar systems for the foreseeable future. Nevertheless, the sheer number of detected extrasolar planets compensates for the coarser measurements.

Searching for Planets

This review focuses on planet populations that are detectable in large numbers by current transit and Doppler surveys: low-mass planets orbiting within about one astronomical unit (AU, the Earth-Sun distance) of their host stars and gas giant planets orbiting within several AU. The Doppler technique has detected and characterized ~700 planets orbiting ~400 stars (*1*, *2*). The Kepler space telescope has discovered more than 2700 planet candidates (*3*–*5*), of which only 5 to 10% are likely to be false-positive detections (*6*, *7*). Giant planets in more distant orbits have also been detected through imaging (*8*) and gravitational microlensing surveys (*9*).

With the Doppler technique, planet masses and orbits are inferred from the observed motions of their host stars. Stellar orbits are point reflections of their planets' orbits, scaled down by the planet-to-star mass ratios. These orbits are measured by the star's line-of-sight velocity (radial velocity, RV) by using high-resolution spectroscopy with ground-based telescopes. Planets are detected by analyzing the repeating patterns in the time series RV measurements and are characterized by their orbital periods (P), minimum masses ($M\sin i$, where M is a planet mass and i is the inclination of a planet's orbit relative to the sky plane), and orbital eccentricities (Fig. 1). Planets with larger masses and shorter orbital periods orbiting lower-mass stars are more detectable. Sensitivity to planets varies from star to star and depends on details of the observing history, including the number, precision, and time baseline of the RV measurements. The earliest Doppler surveys began 20 to 25 years ago with a few hundred nearby, bright stars and are now sensitive to analogs of Jupiter and Saturn. With measurement precisions of ~1 m s^{-1}, more recent surveys are detecting planets of a few Earth masses (M_E) for close-in orbits.

With the transit technique, the eclipses of planets whose orbits happen to be viewed edge-on are detected as brief dips in the host star's brightness (Fig. 1). The size of the planet relative to the star is inferred from the depth of the transit. Jupiter-size planets block ~1% of the flux from Sun-like stars and are detectable with ground-based telescopes, whereas the 0.01%-deep transits of Earth-size planets are only detectable with precise, space-borne telescopes such as the Kepler. The planet's orbital period is the time interval between consecutive transits, and the orbital distance (semi-major axis) can be inferred from Kepler's third law. The mass of a transiting planet can be measured from follow-up Doppler observations if the host star is bright enough and the Doppler amplitude large enough. Masses can also be measured in special cases from precise timing of consecutive transits, which deviate from strict periodicity when multiple planets orbiting the same star gravitationally perturb one another (*10*, *11*).

Transiting planets also offer the opportunity to measure the stellar obliquity—the angle between the stellar spin axis and a planet's orbital axis (*12*)—as well as characteristics of the planet's atmosphere (*13*). Obliquities have primarily been measured as the transiting planet alternately blocks blue-shifted and red-shifted portions of the rotating stellar disk causing apparent Doppler shifts (the Rossiter-McLaughlin effect) (Fig. 1). Obliquity measurements are sensitive to past dynamical interactions that can perturb planets into highly inclined orbits.

Taken together, Doppler and transit detections probe the bulk physical properties of planets (masses, radii, and densities) and their orbital architectures (the number of planets per system and their orbital separations, eccentricities, and geometries). In an observational survey, a large number of stars are searched for planets and the statistical properties of the detected population are studied in order to infer mechanisms of planet formation and evolution. Surveys count planets and naturally produce number distributions of planet parameters (such as the number of detected planets versus planet mass), but these distributions can systematically hide planets that are more difficult to detect. To measure planet occurrence—how commonly planets with a particular characteristic exist in nature—surveys must estimate their sensitivity to planets with different values of that characteristic and statistically correct for their incomplete sample of detected planets.

Abundant Close-In, Low-Mass Planets

Planets intermediate in size between Earth and Neptune are surprisingly common in extrasolar systems, but notably absent from our solar system. The planet size and mass distributions (Fig. 2) show clearly that these small planets outnumber large ones, at least for close-in orbits. Two separate Doppler surveys (*14*, *15*) of nearby, Sun-like stars have shown that planet occurrence rises significantly with decreasing planet mass ($M\sin i$) from 1000 M_E (3 Jupiter masses) down to 3 M_E. In the Eta-Earth Survey, an unbiased set of 166 nearby, G- and K-type stars in the northern sky were observed at Keck Observatory (*14*). The RV of each star was measured dozens of times over 5 years, and the time series RVs were searched for the signatures of planets with orbital periods $P < 50$ days. (The restriction of $P < 50$ days for solar-type stars is equivalent to

Institute for Astronomy, University of Hawai'i at Manoa, 2680 Woodlawn Drive, Honolulu, HI 96822, USA.
E-mail: howard@ifa.hawaii.edu

restricting orbital distances to <0.25 AU, inside of Mercury's 0.4 AU orbit.) In total, 33 planets were detected orbiting 22 of the 166 stars. Low-mass planets ($M\sin i$ = 3 to 30 M_E) were detected more frequently, in spite of their weaker Doppler signals. After correcting for survey incompleteness, the planet mass distribution was fit with a power-law model that rises steeply toward low mass. The probability of a star hosting a close-in planet scales as $(M\sin i)^{-0.48}$. In absolute terms, 15% of Sun-like stars host one or more planets with $M\sin i$ = 3 to 30 M_E orbiting within 0.25 AU, and by extrapolation, another 14% of stars host planets with $M\sin i$ = 1 to 3 M_E.

The HARPS (High Accuracy Radial velocity Planet Searcher) survey measured the RVs of 376 Sun-like stars in the southern sky with somewhat better sensitivity to low-mass planets (*15*). It confirmed the rising planet mass function with decreasing mass and extended it to 1 to 3 M_E planets (Fig. 2). It also demonstrated that low-mass planets have small orbital eccentricities and are commonly found in multi-planet systems with two to four small planets orbiting the same star with orbital periods of weeks or months. The HARPS survey found that at least 50% of stars have one or more planet of any mass with P < 100 days.

The Kepler mission has substantially refined our statistical knowledge of planets between the size of Earth and Neptune by revealing thousands of these planets, compared with the dozens detected with the Doppler technique. The distribution of planet sizes (radii) measured by the Kepler telescope (Fig. 2) follows the same trend as the mass distribution, with small planets being more common (*16*, *17*, *18*). However, the Kepler telescope size distribution extends confidently down to Earth size for close-in planets, whereas the mass distribution is uncertain at the 50% level near 1 Earth mass. The size distribution is characterized by a power-law rise in occurrence with decreasing size (*17*) down to a critical size of ~2.8 Earth radii (R_E), below which planet occurrence plateaus (*16*). Earth-size planets orbiting within 0.25 AU of their host stars are just as common as planets twice that size. The small planets detected by the Kepler telescope (<2 R_E) appear to have more circular orbits than those of larger planets (*19*), suggesting reduced dynamical interactions.

The Kepler telescope is sensitive to sub-Earth-size planets around stars with low photometric noise and has detected planets as small as the Moon (0.3 R_E) (*20*). However, survey sensitivity remains uncertain below 1 R_E with current data, even for orbital periods of a few days. Although the occurrence plateau for 1 to 2.8 R_E with a steep fall-off for larger planets is not well understood theoretically, it offers an important observed property of planets around Sun-like stars that must be reproduced by planet formation models.

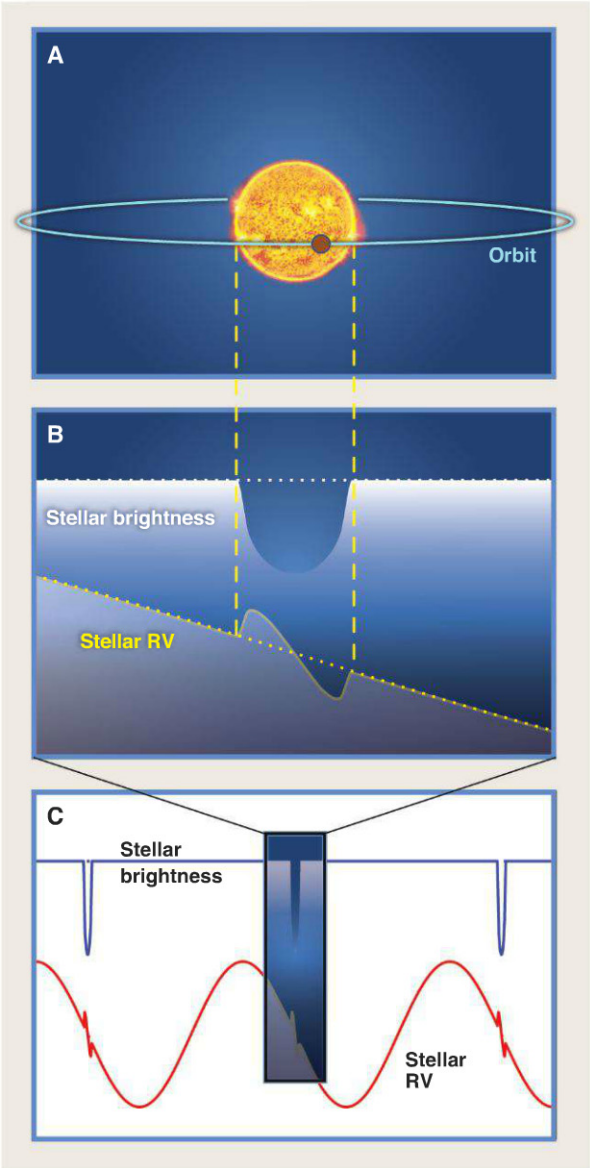


Fig. 1. Schematic illustration of a planetary orbit and the variations in stellar brightness and RV that it causes. (A) A planet orbits its host star and eclipses ("transits") the star as seen by a distant observer. **(B)** Planets are detectable during transit by the decrease in stellar brightness (solid white line). Transit depth is proportional to the blocked fraction of the stellar disk. The stellar obliquity can be measured during transit by anomalous Doppler shifts (the Rossiter-McLaughlin effect; solid yellow line) in the RV time series as the planet blocks portions of the rotating stellar disk. A low-obliquity system with a well-aligned stellar spin axis and planet orbital axis is shown. Nontransiting planets do not produce such effects (dashed lines). **(C)** Over many orbits, planet properties including the size, mass, orbital period, eccentricity, and orbital inclination can be measured from detailed analysis of time-series photometric and RV data.

The high occurrence of small planets with P < 50 days likely extends to more distant orbits. As the Kepler telescope accumulates photometric data, it becomes sensitive to planets with smaller sizes and longer orbital periods. Based on 1.5 years of data, the small-planet occurrence distribution as a function of orbital period is flat out to P = 250 days (with higher uncertainty for larger P). That is, the mean number of planets per star per logarithmic period interval is proportional to $P^{+0.11 \pm 0.05}$ and $P^{-0.10 \pm 0.12}$ for 1 to 2 R_E and 2 to 4 R_E planets, respectively (*21*).

Of the Kepler telescope's planet-host stars, 23% show evidence for two or more transiting planets (*3*). To be detected, planets in multi-transiting systems likely orbit in nearly the same plane, with mutual inclinations of a few degrees at most, as in the solar system. The true number of planets per star (transiting or not) and their mutual inclinations can be estimated from simulated observations of model planetary systems constrained by the number of single, double, triple (and so on) transiting systems detected with the Kepler telescope (*22*, *23*). One model finds an intrinsic multi-planet distribution with 54, 27, 13, 5, and 2% of systems having 1, 2, 3, 4, and 5 planets with P < 200 days, respectively. Most multi-planet systems (85%) have mutual inclinations of less than 3° (*22*). Comparisons of the Kepler telescope and HARPS planetary systems also suggest mutual inclinations of a few degrees (*24*). This high degree of coplanarity is consistent with planets forming in a disk without substantial dynamical perturbations capable of increasing inclinations.

Orbital period ratios in multi-transiting systems provide additional dynamical information. These ratios are largely random (*25*), with a modest excess just outside of period ratios that are consistent with dynamical resonances (ratios of 2:1, 3:2, and so on) and a compensating deficit inside (*26*). The period ratios of adjacent planet pairs demonstrate that >31, >35, and >45% of two-, three-, and four-planet systems, respectively, are dynamically packed; adding a hypothetical planet would gravitationally perturb the system into instability (*27*).

Masses, Radii, and Densities

Although mass and size distributions provide valuable information about the relative occurrence

of planets of different types, it remains challenging to connect the two. Knowing the mass of a planet does not specify its size, and vice versa. A planet the mass of Earth could have a variety of sizes, depending on the composition and the extent of the atmosphere.

This degeneracy can be lifted for ~200 planets with well-measured masses and radii (Fig. 3), most of which are giant planets. The cloud of measurements in Fig. 3 follows a diagonal band from low-mass/small-size to high-mass/large-size. This band of allowable planet mass/size combinations has considerable breadth. Planets less massive than ~30 M_E vary in size by a factor of 4 to 5, and planets larger than ~100 M_E (gas giants) vary in size by a factor of ~2. For the gas giants, the size dispersion at a given mass largely is due to two effects. First, the presence of a massive solid core (or distributed heavy elements) increases a planet's surface gravity, causing it be more compact. Second, planets in tight orbits receive higher stellar flux and are statistically more likely to be inflated relative to the sizes predicted by atmospheric models (the "hydrogen" curve in Fig. 3). Although it is clear that higher stellar flux correlates with giant planet inflation (*28*), it is unclear how the stellar energy is deposited in the planet's interior. Energy deposited in a planet's outer layers is quickly reradiated unless it is somehow circulated to the interior.

Low-mass planets have an even greater variation in size and inferred compositional diversity. The planet Kepler-10b has a mass of 4.6 M_E and a density of 9 g cm^{-3}, indicating a rock/iron composition. With such a high density, this planet likely has little or no atmosphere (*29*). In contrast, the planet Kepler-11e has a density of 0.5 g cm^{-3} and a mass of 8 M_E. A substantial light-element atmosphere (probably hydrogen) is required to explain its mass and radius combination (*30*). The masses and radii of intermediate planets lead to ambiguous conclusions about composition. For example, the bulk physical properties of GJ 1214b (mass 6.5 M_E, radius 2.7 R_E, density 1.9 g cm^{-3}) are consistent with a several compositions: a "super-Earth," with a rock/iron core surrounded by ~3% H_2 gas by mass; a "water world" planet, consisting of a rock/iron core, a water ocean, and atmosphere that contribute ~50% of the mass; or a "mini-Neptune," composed of rock/iron, water, and H/He gas (*31*). For this particular planet, measuring the transmission spectrum during transit appears to have lifted the degeneracy. The small atmospheric scale height of GJ1214b favors a high mean molecular weight atmosphere (possibly water) but is also consistent with an H_2 or H/He atmosphere rendered featureless by thick clouds (*32*).

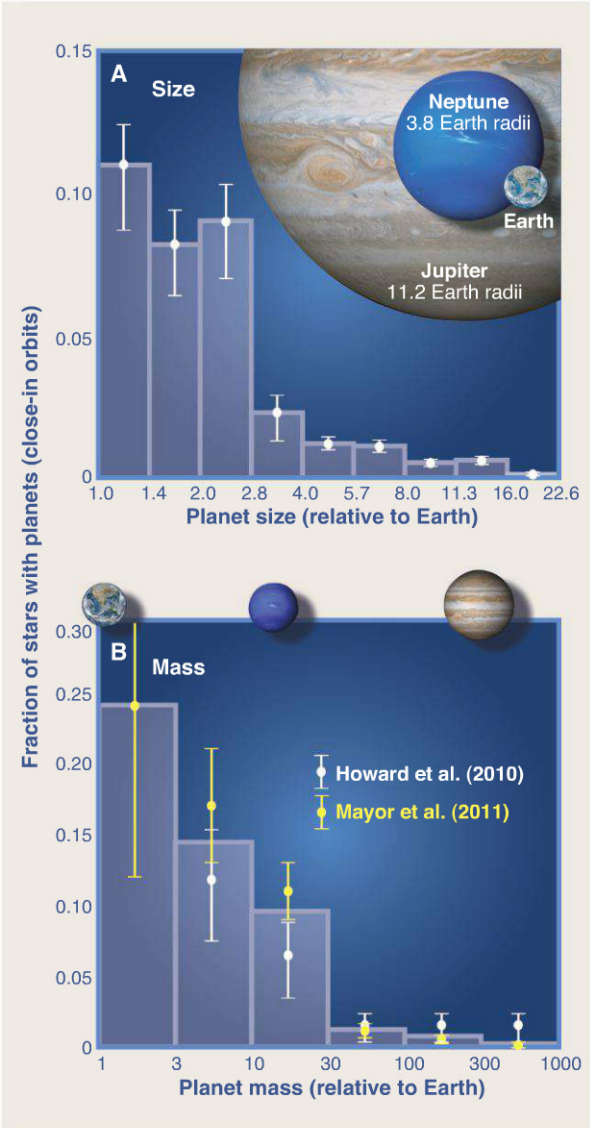


Fig. 2. The (A) size and (B) mass distributions of planets orbiting close to G- and K-type stars. The distributions rise substantially with decreasing size and mass, indicating that small planets are more common than large ones. Planets smaller than 2.8 R_E or less massive than 30 M_E are found within 0.25 AU of 30 to 50% of Sun-like stars. (A) The size distribution from transiting planets shows occurrence versus planet radius and is drawn from two studies of Kepler telescope data: (*16*) for planets smaller than four times Earth's size and (*17*, *59*) for larger planets. The inset images of Jupiter, Neptune, and Earth show their relative sizes. The mass (*M*sin*i*) distributions (B) show the fraction of stars having at least one planet with an orbital period shorter than 50 days (orbiting inside of ~0.25 AU) are from the Doppler surveys from (*14*) (white points) and (*15*) (yellow points), whereas the histogram shows their average values. Inset images of Earth, Neptune, and Jupiter are shown on the top horizontal axis at their respective masses. Both distributions are corrected for survey incompleteness for small/low-mass planets to show the true occurrence of planets in nature.

Gas Giant Planets

The orbits of giant planets are the easiest to detect by using the Doppler technique and were the first to be studied statistically (*33*, *34*). Observations over a decade of a volume-limited sample of ~1000 F-, G-, and K-type dwarf stars at Keck Observatory showed that 10.5% of G- and K-type dwarf stars host one or more giant planets (0.3 to 10 Jupiter masses) with orbital periods of 2 to 2000 days (orbital distances of ~0.03 to 3 AU). Within those parameter ranges, less massive and more distant giant planets are more common. A fit to the giant planet distribution in the mass-period plane shows that occurrence varies as $M^{-0.31 \pm 0.2}P^{+0.26 \pm 0.1}$ per logarithmic interval $d\log M\, d\log P$. Extrapolation of this model suggests that 17 to 20% of such stars have giant planets orbiting within 20 AU (P = 90 years) (*35*). This extrapolation is consistent with a measurement of giant planet occurrence beyond ~2 AU from microlensing surveys (*36*). However, the relatively few planet detections from direct imaging planet searches suggest that the extrapolation is not valid beyond ~65 AU (*37*).

These overall trends in giant planet occurrence mask local pile-ups in the distribution of orbital parameters (*38*). For example, the number distribution of orbital distances for giant planets shows a preference for orbits larger than ~1 AU and to a lesser extent near 0.05 AU, where "hot Jupiters" orbit only a few stellar radii from their host stars (Fig. 4A). This "period valley" for apparently single planets is interpreted as a transition region between two categories of planets with different migration histories (*33*). The excess of planets starting at ~1 AU approximately coincides with the location of the ice line. Water is condensed for orbits outside of the ice line, providing an additional reservoir of solids that may speed the formation of planet cores or act as a migration trap for planets formed farther out (*39*). The orbital period distribution for giant planets in multi-planet systems is more uniform, with hot Jupiters nearly absent and a suppressed peak of planets in >1 AU orbits. The giant-planet eccentricity distribution (Fig. 4B) also differs between single- and multi-planet systems. The eccentricity distribution for single planets can be reproduced quantitatively with a dynamical model in which initially low eccentricities are excited by planet-planet scattering

(*40*). Multi-planet systems likely experienced substantially fewer scattering events. One interpretation is that eccentric, single-planet systems are the survivors of scattering events that ejected the other planets in the system.

Although hot Jupiters (giant planets with $P <$ ~10 days) are found around only 0.5 to 1.0% of Sun-like stars (*41*), they are the most well-characterized planets because they are easy to detect and follow up with ground- and space-based telescopes. However, they have a number of unusual characteristics and a puzzling origin. In contrast to close-in small planets, hot Jupiters are commonly the only detected planet orbiting the host star within observational limits (Fig. 4) (*42*). Many have low eccentricities, primarily because of tidal circularization. The measured obliquities of stars hosting hot Jupiters display a peculiar pattern: Obliquities are apparently random above a critical stellar temperature of ~6250 K, but cooler systems are mostly aligned. In situ formation is unlikely for hot Jupiters because of insufficient protoplanetary disk mass so close to the star. Rather, they likely formed in the disk at several AU, were gravitationally perturbed into orbits with random inclinations and high eccentricities, and were captured at ~0.05 AU by dissipation of orbital energy in tides raised on the planet. For systems with sufficiently strong tides raised by the planet on the star (which depend on a stellar convective zone that is only present below ~6250 K), the stellar spin axis aligns to the orbital axis (*12*).

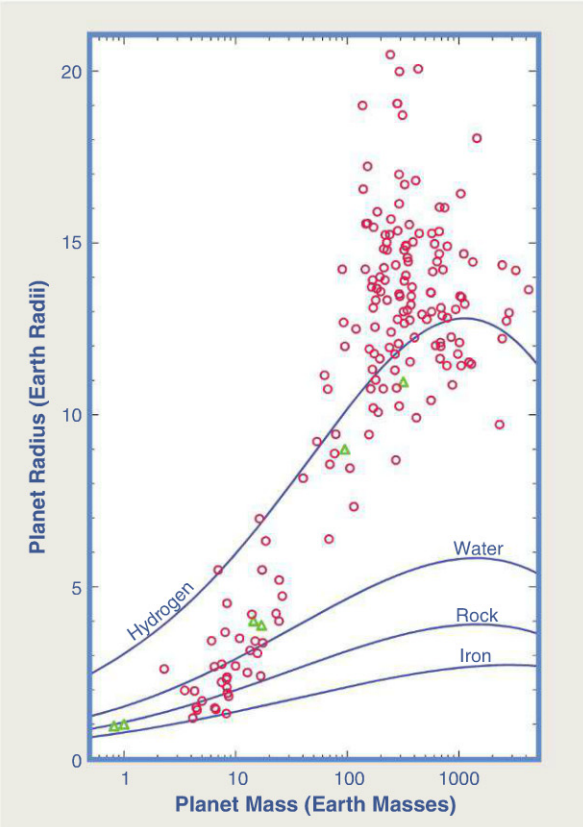


Fig. 3. Masses and sizes of well-characterized planets. Extrasolar planets (*1*, *58*, *60*) are shown as open red circles, whereas solar system planets are designated by open green triangles. Radii were measured by means of transit photometry, and masses were measured by radial velocity or transit timing methods. Model mass-radius relationships for idealized planets consisting of pure hydrogen (*61*), water, rock (Mg_2SiO_4), or iron (*62*) are shown as blue lines. Poorly understood heating mechanisms inflate some gas giant planets (larger than ~8 R_E) to sizes larger than predicted by the simple hydrogen model. Smaller planets (less massive than ~30 M_E) show great diversity in size at a fixed mass, likely because of varying density of solids and atmospheric extent. Gas giant planets are overrepresented relative to their occurrence in nature due to their relative ease of detection and characterization.

Planet Formation

Metal-rich stars (*43*) are more likely to host giant planets within 5 AU. This "planet-metallicity correlation" was suggested in 1997, when the first four Doppler-discovered giant planets were all found to orbit stars with higher iron abundance than that of the Sun (*44*). Initially, this correlation was seen as an artifact of stellar self-pollution from accretion of metals onto the stellar atmosphere during planet formation. Today, it provides evidence for the core accretion model of planet formation (*45*). In this model, a high density of solids in the protoplanetary disk is required to form giant planets, which pass through two key phases. Protoplanets must grow to masses of ~5 to 10 M_E by accretion of solids (dust and ice) from the disk. The protoplanet then undergoes runaway gas accretion, increasing its mass by an order of magnitude, but only if the protoplanetary gas has not yet dissipated. Giant planet formation is a race against dispersal of gas from the protoplanetary disk on a time scale of ~3 to 5 million years (*46*). This theory predicts that giant planets should be more common around massive and metal-rich stars whose disks have higher surface densities of solids.

The planet-metallicity correlation was validated statistically by Doppler surveys of ~1000 stars with masses of 0.7 to 1.2 solar masses (M_{Sun}) and uniformly measured metallicities, which were highly sensitive to giant planets (*47*, *48*). The probability of a star hosting a giant planet is proportional to the square of the number of iron atoms in the star relative to that in the Sun, $p(\text{planet}) \propto N_{Fe}^2$ (*47*). A later Doppler study spanned a wider range of stellar masses (0.3 to 2.0 M_{Sun}) and showed that the probability of a star hosting a giant planet correlates with both stellar metal content and stellar mass, $p(\text{planet}) \propto N_{Fe}^{1.2 \pm 0.2} M_{star}^{1.0 \pm 0.3}$ (*49*). The planet-metallicity correlation only applies to gas giant planets. Planets larger than 4 R_E (Neptune size) preferentially orbit metal-rich stars, whereas smaller planets are found in equal numbers around stars with a broad range of metallicities (*50*). That is, small planets form commonly in most protoplanetary disks, but only a fraction grow to a critical size in time to become gas giants.

Although the planet-metallicity trends support the basic mechanism of the core accretion model, many statistical features of the observed planet population cannot yet be explained in detail. In particular, the population of low-mass planets inside of 1 AU is difficult to reproduce in conventional models. Population synthesis models attempt to follow the growth and migration of sub-Earth-size protoplanets in a protoplanetary disk to predict the planet masses and orbital distances after the disk dissipates (*39*, *51*). These models reproduce the giant planet population well, but struggle with low-mass planets. In population synthesis models, low-mass planets form primarily near and beyond the ice line (several AU) and migrate to close orbital distances through interactions with the disk. The prescriptions for migration and growth in these models produce "deserts" of reduced planet occurrence precisely where Doppler and transit surveys detect a great abundance of planets.

An alternative model is in situ formation of close-in, low-mass planets with minimal subsequent planet migration (*52*, *53*). Although this model correctly reproduces several observed properties of close-in planets (the mass distribution, multi-planet frequencies, and small eccentricities and inclinations), it is still in the early stages of development. One challenge is that in situ formation requires ~20 to 50 M_E of protoplanetary disk material inside of 1 AU, which is poorly constrained by observations.

Earth-Size Planets in the Habitable Zone

The detection of planets the size or mass of Earth remains a prominent observational goal. Using the Doppler technique, one such planet has been detected: a planet with $M\sin i = 1.1\ M_E$ orbiting the nearby star α Centauri B with an orbital separation of 0.04 AU that renders it too hot for life (*54*). The Doppler signal from an Earth-mass planet orbiting at 1 AU is smaller by a factor of five, beyond the reach of current instrumentation and possibly hidden in

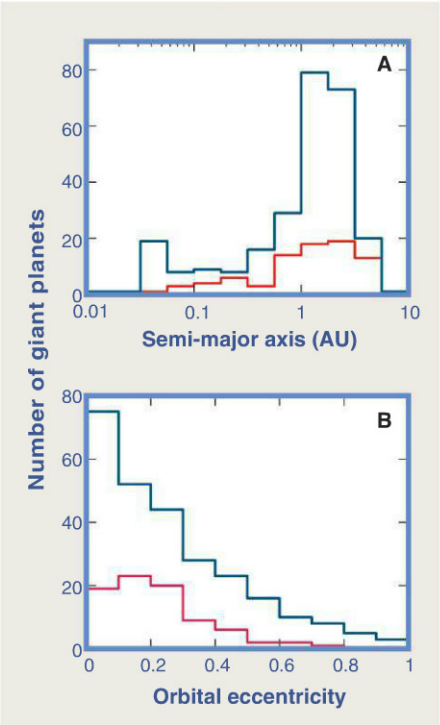


Fig. 4. Orbital characteristics of giant planets (*M*sin*i* > 0.2 Jupiter masses) detected by Doppler surveys (*1*, *60*). (**A**) The number distribution of semi-major axes shows that apparently single planets (blue) preferentially orbit at distances of ~0.05 AU and at ~1 to 3 AU from their host stars. These preferred orbits are diminished in multi-planet systems (red). The decline in number of detected planets for orbits outside of ~3 AU is not significant; fewer stars have been searched for such planets as compared with the closer orbits. (**B**) The distribution of orbital eccentricities for apparently single planets (blue) span the full range, with low-eccentricity orbits being more common. Giant planets in multi-planet systems (red) have orbits that are more commonly close to circular (eccentricity = 0). The larger eccentricities of single planets suggests that they were dynamically excited from a quiescent, nearly circular origin, perhaps by planet-planet scattering that resulted in the ejection of all but one detectable planet per system.

Doppler noise from the star. Nevertheless, Doppler planet searches continue. If a planet the size or mass of the Earth is detected orbiting in the habitable zone of a nearby star, it would be a milestone for science and could catalyze the development of instruments to image and take spectra of such planets.

The Kepler telescope has detected dozens of Earth-size planets, although these planets orbit interior to their stars' habitable zones (*16*, *18*). The habitable zone—the set of planetary orbits consistent with liquid water existing on the planet's surface—is challenging to define precisely because it depends on the detailed energy balance for a planet with an often poorly constrained composition and atmosphere (*55*). Nevertheless, the Kepler telescope has also detected planets slightly larger than Earth (1.4 and 1.6 times Earth size) in the classically defined habitable zone (*56*). The primary goal for the Kepler telescope in its extended mission is to detect individual Earth-size planets in the habitable zone and to estimate their occurrence rate. Not all of the Earth-size planets detected by the Kepler telescope will be 1 M_E, however. Measuring the densities of several Earth-size planets (not necessarily in the habitable zone) will offer some constraints on the typical compositions of Earth-size planets.

Targeting low-mass stars offers a shortcut in the search for planets the size and mass of Earth. Planets are more detectable in Doppler and transit searches of lower-mass stars. The habitable zones around such stars are also closer, owing to the reduced brightness of low-mass stars. Small planets may be more common around low-mass stars as well (*17*) [an opposing view is available in (*18*)]. An analysis of planets discovered with the Kepler telescope orbiting M dwarfs suggests a high rate of overall planet occurrence, 0.9 planets per star in the size range 0.5 to 4 R_E in $P <$ 50 day orbits. Earth-size planets (0.5 to 1.4 R_E) are found in the habitable zones of 15^{+13}_{-6}% of the M dwarfs in the Kepler telescope sample (*57*). As the Kepler telescope's sensitivity expands during the extended mission, we will likely learn how common Earth-size planets are in the habitable zones of Sun-like stars.

References and Notes

1. J. T. Wright *et al.*, *Proc. Astron. Soc. Pacific* **123**, 412 (2011).
2. Updates to (*1*) are available at http://exoplanets.org.
3. C. J. Burke *et al.*, American Astronomical Society Meeting #221, abstract #216.02 (2013).
4. The current catalog of planets discovered with the Kepler telescope are available at http://exoplanetarchive.ipac.caltech.edu/cgi-bin/ExoTables/nph-exotbls?dataset=cumulative.
5. N. M. Batalha *et al.*, *Astrophys. J.* **204**, 24 (2013).
6. T. Morton, J. J. Johnson, *Astrophys. J.* **738**, 170 (2011).
7. The false positive probabilities (FPPs) for Kepler telescope planet candidates are typically small because of detailed vetting of the photometry, including analysis of the transit shape and astrometric displacement. FPPs depend on the host star brightness and crowding in the field as well as on planet size; giant planets often have larger FPPs.
8. A. Vigan *et al.*, *Astron. Astrophys.* **544**, A9 (2012).
9. B. S. Gaudi, *Annu. Rev. Astron. Astrophys.* **50**, 411 (2012).
10. E. Agol, J. Steffen, R. Sari, W. Clarkson, *Mon. Not. R. Astron. Soc.* **359**, 567 (2005).
11. M. J. Holman, N. W. Murray, *Science* **307**, 1288 (2005).
12. S. Albrecht *et al.*, *Astrophys. J.* **757**, 18 (2012).
13. S. Seager, D. Deming, *Annu. Rev. Astron. Astrophys.* **48**, 631 (2010).
14. A. W. Howard *et al.*, *Science* **330**, 653 (2010).
15. M. Mayor *et al.*, *Astron. Astrophys.*, http://arxiv.org/abs/1109.2497 (2011).
16. E. Petigura *et al.*, *Astrophys. J.*, http://arxiv.org/abs/1304.0460 (2013).
17. A. W. Howard *et al.*, *Astrophys. J.* **201**, 15 (2012).
18. F. Fressin *et al.*, *Astrophys. J.*, http://arxiv.org/abs/1301.0842 (2013).
19. P. Plavchan *et al.*, *Astrophys. J. Lett.*, http://arxiv.org/abs/1203.1887 (2012).
20. T. Barclay *et al.*, *Nature* **494**, 452 (2013).
21. S. Dong, Z. Zhu, http://arxiv.org/abs/1212.4853 (2012).
22. J. Fang, J.-L. Margot, *Astrophys. J.* **761**, 92 (2012).
23. J. J. Lissauer *et al.*, *Astrophys. J.* **197**, 8 (2011).
24. P. Figueira *et al.*, *Astron. Astrophys.* **541**, A139 (2012).
25. D. C. Fabrycky *et al.*, *Astrophys. J.*, http://arxiv.org/abs/1202.6328 (2012).
26. Y. Lithwick, Y. Wu, *Astrophys. J.* **756**, L11 (2012).
27. J. Fang, J.-L. Margot, *Astrophys. J.*, http://arxiv.org/abs/1302.7190 (2013).
28. L. M. Weiss *et al.*, *Astrophys. J.*, http://xxx.lanl.gov/abs/1303.2150 (2013).
29. N. M. Batalha *et al.*, *Astrophys. J.* **729**, 27 (2011).
30. J. J. Lissauer *et al.*, *Nature* **470**, 53 (2011).
31. L. A. Rogers, S. Seager, *Astrophys. J.* **716**, 1208 (2010).
32. Z. K. Berta *et al.*, *Astrophys. J.* **747**, 35 (2012).
33. S. Udry, M. Mayor, N. C. Santos, *Astron. Astrophys.* **407**, 369 (2003).
34. G. Marcy *et al.*, *Progr. Theor. Phys. Suppl.* **158**, 24 (2005).
35. A. Cumming *et al.*, *Publ. Astron. Soc. Pac.* **120**, 531 (2008).
36. A. Gould *et al.*, *Astrophys. J.* **720**, 1073 (2010).
37. E. L. Nielsen, L. M. Close, *Astrophys. J.* **717**, 878 (2010).
38. J. T. Wright *et al.*, *Astrophys. J.* **693**, 1084 (2009).
39. S. Ida, D. N. C. Lin, *Astrophys. J.* **685**, 584 (2008).
40. S. Chatterjee, E. B. Ford, S. Matsumura, F. A. Rasio, *Astrophys. J.* **686**, 580 (2008).
41. J. T. Wright *et al.*, *Astrophys. J.* **753**, 160 (2012).
42. J. H. Steffen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 7982 (2012).
43. All elements except for hydrogen and helium are considered "metals" by astronomers. "Metallicity" refers to the abundance of metals (often iron) relative to solar proportions.
44. G. Gonzalez, *Mon. Not. R. Astron. Soc.* **285**, 403 (1997).
45. Self-pollution was ruled out because evolved subgiant planet–host stars are as metal-rich as are main sequence–planet hosts, despite convective zone mixing in evolved stars that would have diluted their surface-layer pollution signature (*53*).
46. M. C. Wyatt *et al.*, *Annu. Rev. Astron. Astrophys.* **46**, 339 (2008).
47. D. A. Fischer, J. Valenti, *Astrophys. J.* **622**, 1102 (2005).
48. N. Santos, G. Israelian, M. Mayor, *Astron. Astrophys.* **415**, 1153 (2004).
49. J. A. Johnson, K. M. Aller, A. W. Howard, J. R. Crepp, *Proc. Astron. Soc. Pacific* **122**, 905 (2010).
50. L. A. Buchave *et al.*, *Nature* **486**, 375 (2012).
51. C. Mordasini, Y. Alibert, W. Benz, *Astron. Astrophys.* **501**, 1139 (2009).
52. B. M. S. Hansen, N. Murray, *Astrophys. J.* **751**, 158 (2012).
53. E. Chiang, G. Laughlin, *Monthly Notices of the Royal Astronomical Society*, http://arxiv.org/abs/1211.1673 (2012).
54. X. Dumusque *et al.*, *Nature* **491**, 207 (2012).
55. S. Seager, *Science* **340**, 577 (2013).
56. W. J. Borucki *et al.*, *Science* **340**, 587 (2013).
57. C. D. Dressing *et al.*, *Astrophys. J.*, http://arxiv.org/abs/1302.1647 (2013).
58. G. Marcy *et al.*, American Astronomical Society Meeting #221, abstract #231.01 (2013)
59. For planets smaller than 4 R_E, I adopt Kepler telescope occurrence measurements for orbital periods of 5 to 50 days, which include corrections for missed planet detections in the Kepler telescope pipeline (*16*). For larger planets, I use measurements of a larger number of stars, but for all orbital periods <50 days (*17*).
60. The properties of extrasolar planets plotted in Figs. 3 and 4 were downloaded from the Exoplanet Orbit Database (http://exoplanets.org) on 14 February 2013. Additionally, *Kepler*-discovered planets with ≥2-σ mass measurements from (*58*) were included for Fig. 3.
61. S. Seager, M. Kuchner, C. A. Hier-Majumder, B. Militzer, *Astrophys. J.* **669**, 1279 (2007).
62. J. J. Fortney, M. S. Marley, J. W. Barnes, *Astrophys. J.* **659**, 1661 (2007).

Acknowledgments: I thank K. Teramura for assistance preparing the graphics and G. Marcy, D. Fischer, J. Wright, T. Currie, E. Petigura, H. Isaacson, R. Dawson, I. Crossfield, S. Kane, J. Steffen, J. Maurer, and S. Howard for comments on this manuscript. I also thank three anonymous reviewers for constructive criticism. I acknowledge partial funding from NASA grant NNX12AJ23G.

10.1126/science.1233545

REVIEW

Exoplanet Habitability

Sara Seager

The search for exoplanets includes the promise to eventually find and identify habitable worlds. The thousands of known exoplanets and planet candidates are extremely diverse in terms of their masses or sizes, orbits, and host star type. The diversity extends to new kinds of planets, which are very common yet have no solar system counterparts. Even with the requirement that a planet's surface temperature must be compatible with liquid water (because all life on Earth requires liquid water), a new emerging view is that planets very different from Earth may have the right conditions for life. The broadened possibilities will increase the future chances of discovering an inhabited world.

For thousands of years people have wondered, "Are we alone?" Now, for the first time in human history, the answer to this and other long-standing questions in the search for life beyond our solar system may finally be in reach through the observation and study of exoplanets—planets orbiting stars other than the Sun.

The research field of exoplanets has grown dramatically since the first planet orbiting a Sun-like star was discovered nearly 20 years ago (*1*). Nearly 1000 exoplanets are known to orbit nearby stars, a few thousand more planet "candidates" have been identified, and planets are so common that on average every star in the Milky Way should have at least one planet (*2*, *3*). The numbers of exoplanet candidates found by NASA's Kepler space telescope are high enough that robust statements of the frequency of their occurrence is possible, including the astonishing finding that small planets by far outnumber large planets in our galaxy (*3*, *4*), and the first statement about how common Earth-size planets are in the habitable zones of small stars (*5*).

The habitable zone is a region around a star where a planet can have surface temperatures consistent with the presence of liquid water. All life on Earth requires liquid water, so the planetary surface-temperature requirement appears to be a natural one. The climates of planets with thin atmospheres are dominated by external energy input from the host star, so that a star's "habitable zone" is based on distance from the host star. Small stars have a habitable zone much closer to them as compared to Sun-like stars, owing to their lower luminosity. The habitable zone was first discussed in the mid-20th century, inspired by attempts to understand the climate of early Earth and Mars (*6*, *7*), and was later brought onto a self-consistent footing when the carbonate-silicate cycle was proposed as a climate-stabilizing mechanism (*8*, *9*).

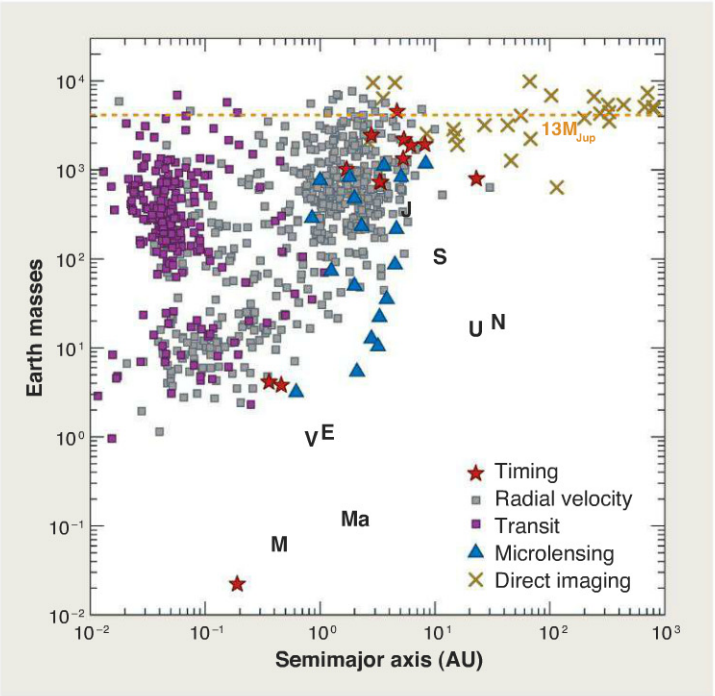


Fig. 1. Known exoplanets as of March 2013. Exoplanets are found at a nearly continuous range of masses and semimajor axes. Many different techniques are successful at discovering exoplanets, as indicated by the different symbols. The solar system planets are denoted by the first one or two letters of their name. The horizontal line is the conventional upper limit to a planet mass, 13 Jupiter masses. The sloped, lower boundary to the collection of gray squares is due to a selection effect in the radial velocity technique. Small planets are beneath the threshold for the current state of almost all exoplanet detection techniques. Data are from http://exoplanet.eu/.

The habitable zone for exoplanets was first presented and modeled in detail by (*9*), who also suggested an empirical version based on the concept that both Venus [0.7 astronomical units (AU) from the Sun, where an AU is the Earth-Sun distance] and Mars (1.5 AU) may have had liquid surface water at some point in the past. Most exoplanet habitable-zone research that followed continued to focus on terrestrial-like planet atmospheres orbiting main-sequence stars [see (*10*) and references therein]. This article reviews updates to the habitable zone and their rationale.

A planet in the habitable zone has no guarantee of actually being habitable. Venus and Earth may both be argued as being in the Sun's habitable zone and would appear from exoplanet discovery techniques to be the same size and mass. Yet, Venus is completely hostile to life owing to a strong greenhouse effect and resulting high surface temperatures (>700 K), whereas Earth has the right surface temperature for liquid water oceans and is teeming with life.

If there is one important lesson from exoplanets, it is that anything is possible within the laws of physics and chemistry. Planets of almost all masses, sizes, and orbits have been detected (Fig. 1), illustrating not only the stochastic nature of planet formation but also a subsequent migration through the planetary disk from the planet's place of origin [e.g., (*11*)]. The huge diversity of exoplanets and the related anticipated variation in their atmospheres, in terms of mass and composition, have motivated a strong desire to revise the view of planetary habitability. In parallel, there is a growing acceptance that even in the future, the number of suitable planets accessible to detailed follow-up observations may be very small. To maximize our chances of identifying a habitable world, a broader understanding of which planets are habitable is a necessity.

Habitable Planets, Conventionally Defined

The conventionally habitable planet is one with surface liquid water. Water is required for all life as we know it, and has motivated a mantra in astrobiology, "follow the water." Challenging the water requirement paradigm, a National Academies report (*12*) concluded that although a liquid environment is required by life, it need not be limited to water. In the search for life beyond the solar system, however, we still focus on environments that support liquid water simply because

Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

E-mail: seager@mit.edu

water is the most accessible, abundant, and common liquid in terms of planetary material (*13*).

For illustration and review, we consider water on the terrestrial planets in our own solar system. Earth is touted as the "Goldilocks planet"—not too hot, not too cold, but just right for surface liquid water (*14*). Venus, 30% closer to the Sun than Earth and receiving 90% more radiation from the Sun, may have had liquid water oceans billions of years ago, as possibly implied by the elevated deuterium/hydrogen (D/H) ratio in the venusian atmosphere (*15*). Because of warm surface temperatures, water evaporated to saturate the upper atmosphere where solar extreme ultraviolet (EUV) radiation photodissociated the H_2O, enabling H to escape to space. The increasing atmospheric water vapor further warmed the surface, creating a positive feedback loop that led to a "runaway greenhouse effect," which caused Venus to rapidly lose its oceans [but compare (*16*)]. Mars, at 1.5 AU from the Sun, is thought to have had at least episodic surface liquid water in the past, based predominantly on geomorphological features [e.g., (*17*)]. Mars was too small to hold onto a warming atmosphere and is now so cold there is no place on the Martian surface where water could be liquid. The habitable zone for terrestrial-type exoplanets with terrestrial-like atmospheres of various masses and CO_2 concentration are described in (*10*) and result in a habitable zone of 0.99 to 1.7 AU (Fig. 2). The inner edge of the habitable zone is determined by loss of water via the runaway greenhouse effect (*18*) and the outer edge by CO_2 condensation.

For exoplanets, we cannot directly observe liquid surface water (*19*). Atmospheric water vapor may be used as a proxy; as long as a temperate planet is small or of low enough mass, water vapor should not be present because water will be photodissociated with H escaping to space. Atmospheric water vapor has been detected on hot giant transiting exoplanets [e.g., (*20*)] and is highly sought after for the mini Neptune GJ 1214b [e.g., (*21*)]. Both of these types of planets are too hot for surface liquid water [for a discussion of GJ 1214b, see (*22*)]; notably, water vapor will be naturally occurring on planets that are massive enough or cold enough to hold on to water vapor molecules. The detection of water vapor in the atmosphere of smaller, more terrestrial-like planets is currently out of reach.

Given the observational inaccessibility of the key habitability indicator water vapor on terrestrial-like exoplanets, the habitable zone around a star is a powerful guide for astronomers because it tells us where to focus future efforts of exoplanet discovery. We must redefine the habitable-zone concept, however, given the expected and observed diversity of exoplanets.

The Diversity of Exoplanets and the Controlling Factors of Habitability

Taking surface liquid water as a requirement, what types of planets are habitable? Water is in the liquid phase for a range of temperatures and pressures. Planets should also have a wide range of surface temperatures and pressures, expected from their diversity in mass and size and likely atmospheres. If we could connect the liquid water phase diagram with planet surface conditions, broadly speaking, we would know to first order which planets may be habitable.

The water phase diagram can be used as a qualitative guide to show that pressures thousands of times higher than Earth's 1-bar surface pressure can maintain liquid water at high temperatures (*23*). A suitable temperature for life can be considered to be between the freezing point of water and the upper temperature limits for life, about 395 K (*24*). A notable inaccuracy in the phase diagram is that the water phase boundaries at high pressures have not been studied for a variety of gas mixtures relevant for exoplanets (*25*).

The surface temperature on an exoplanet is governed by the atmosphere's greenhouse gases (or lack thereof). Specifically, the greenhouse gases absorb and reradiate energy from the host star, in the form of upwelling infrared (IR) radiation from the planet's surface. Whereas on Earth we are concerned with, e.g., parts-per-million rise in the greenhouse gas CO_2 concentrations, for potentially habitable exoplanets we do not know a priori and cannot yet measure what gases are in the atmosphere even to the tens of percent level. The atmospheric mass and composition of any specific small exoplanet is not predictable (*26*).

Nevertheless, it is worth summarizing some key factors controlling a planet atmosphere's greenhouse gas inventory. A planet's atmosphere forms from outgassing during planet formation or is gravitationally captured from the surrounding proto-planetary nebula. For terrestrial planets, the primordial atmosphere may be completely changed by escape of light gases to space, continuous outgassing from an active young interior, and bombardment by asteroids and comets. At a later stage, the physical processes operating at the top or bottom of the atmosphere still sculpt the atmosphere. These physical processes are well studied by exoplanet theorists but often with controversy or no conclusion. For example, atmospheric escape is induced by the host star's EUV flux and carried out by a number of thermal or nonthermal escape mechanisms. But the star's past EUV flux, which of the escape mechanisms was at play, and whether or not the planet has a protective magnetic field are not known [e.g., (*27*)]. As a second example, at the bottom of the atmosphere, plate tectonics and volcanic outgassing contribute to burial and recycling of atmospheric gases, but arguments as to whether or not plate tectonics will occur in a super-Earth planet more massive than Earth are still under debate (*28*, *29*). A long list of other surface and interior processes affect the atmospheric composition, including but not limited to the ocean fraction for dissolution of CO_2 and for atmospheric relative humidity, redox state of the planetary surface and interior, acidity levels of the oceans planetary albedo, and surface gravity [for

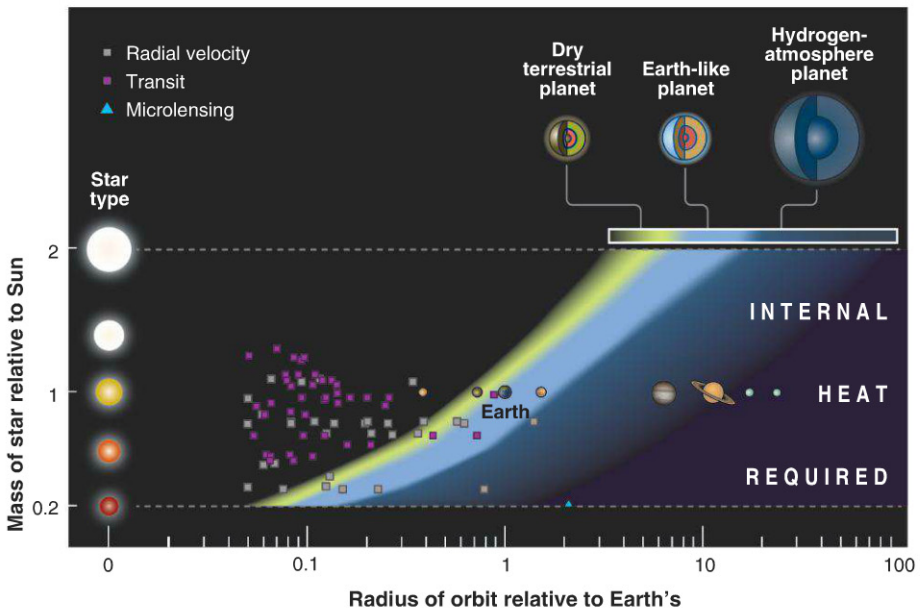


Fig. 2. The habitable zone. The light blue region depicts the "conventional" habitable zone for planets with N_2-CO_2-H_2O atmospheres (*9*, *10*). The yellow region shows the habitable zone as extended inward for dry planets (*36*, *37*), as dry as 1% relative humidity (*37*). The outer darker blue region shows the outer extension of the habitable zone for hydrogen-rich atmospheres (*34*) and can extend even out to free-floating planets with no host star (*35*). The solar system planets are shown with images. Known exoplanets are shown with symbols [here, planets with a mass or minimum mass less than 10 Earth masses or a radius less than 2.5 Earth radii taken from (*66*)].

a more detailed list, see (*30*)]. Many other factors are relevant to habitability, including the radiation environment from the star, especially the energy distribution as a function of wavelength and the EUV radiation that destroys molecules and determines their atmospheric lifetime, and x-ray fluxes that could be detrimental to surface life (*27*). In some cases, planets have been found to orbit one or both stars of a binary star system, complicating the influence of stellar radiation.

A Major Extension of the Habitable Zone

For our qualitative assessment of habitability, we therefore focus on the dominant planetary atmospheric greenhouse gases and how they delimit the habitable zone (Fig. 2).

The most important atmospheric greenhouse gas that extends the habitable zone to large planet-star separation is molecular hydrogen (H_2). Planets are expected to form with some primordial light gases, either H_2 (from interior outgassing) (*26*, *31*) or H_2 and He (from gravitational capture of gas from the surrounding protoplanetary disk). Although small planets like Earth, Venus, and Mars are unable to retain these light gases, more massive or colder exoplanets are expected to be able to do so. H_2 is a formidable greenhouse gas, because it can absorb radiation over a wide, continuous wavelength range. Most molecules absorb in discreet bands. As a homonuclear molecule, H_2 does not have a dipole moment and therefore lacks the typical rotational-vibrational bands that absorb light at near-IR wavelengths. However, a momentary dipole is induced by collisions, and thus at high enough pressures, frequent collisions induce very broadband absorption (*32*, *33*). Furthermore, H_2 does not condense until tens of kelvin at 1- to 100-bar pressures (in comparison, CO_2 condenses at about 190 to 250 K for 1- to 10-bar pressures and is therefore a cutoff for the cold end of conventional planet habitability). The potency of H_2 as a greenhouse gas means that planets can have surface liquid water at a factor of several times larger planet-sun separations than planets with CO_2 atmospheres (*34*) and even possibly extending to rogue planets that were ejected from their birth planetary system and are now floating through the galaxy (*35*).

The inner edge of the habitable zone is controlled by the strong greenhouse gas H_2O, which is fundamentally unavoidable on habitable worlds. Surface liquid water—the adopted requirement for habitability—gives rise to atmospheric water vapor. The habitable planets closest to their host stars must therefore be relatively dry (*36*, *37*)—that is, with a smaller ocean-land fraction than Earth—so the atmosphere will have less water vapor than Earth's. But the putative inner-edge habitable-zone planet must not be too dry; otherwise, CO_2 cannot be washed out of the atmosphere, which would lead to a buildup of CO_2 and subsequent warming. Theoretical simulations of planet formation indicate that dry planets are possible [e.g., (*38*)].

Pockets of small areas of habitability on an individual exoplanet are usually disregarded for exoplanets (*39*); the concern is that they will not lead to any detectable atmospheric signatures. Although large planetary moons of giant planets may be habitable (some might even have interior energy generated by planet-moon tidal friction), detectability of the moon's atmosphere is a concern because of severe contamination from the adjacent larger, brighter planet.

We have seen that planetary habitability is very planet-specific. The habitable zone has been defined with an inner edge of about 0.5 AU around a solar-like star, for a dry rocky planet (*37*), out to 10 AU around a solar-like star for a planet with an H_2 atmosphere and no interior energy (*34*), and even possibly out to free-floating planets with no host star, for planets with thick H_2 atmospheres (*35*) (Fig. 2).

Ideally, we would triage each planet first by the planet's bulk density, using a measured mass and radius, to screen planets for those that have thin atmospheres. Next, we could use the star's luminosity and planet-star separation, as well as model possibilities of the planet's interior, to assess whether the likely surface temperatures are conducive to support liquid water. For planets that pass the tests, telescopic observations of the planet's atmosphere to identify water vapor as a proxy for surface liquid water would be a definitive step for identifying a habitable world. However, for most exoplanets, such fundamental measurements will not be possible. In some cases, a planet's mass but not size can be measured; in other cases, the size but not mass can be measured, and the atmosphere will be accessible only for a few small planets orbiting nearby stars.

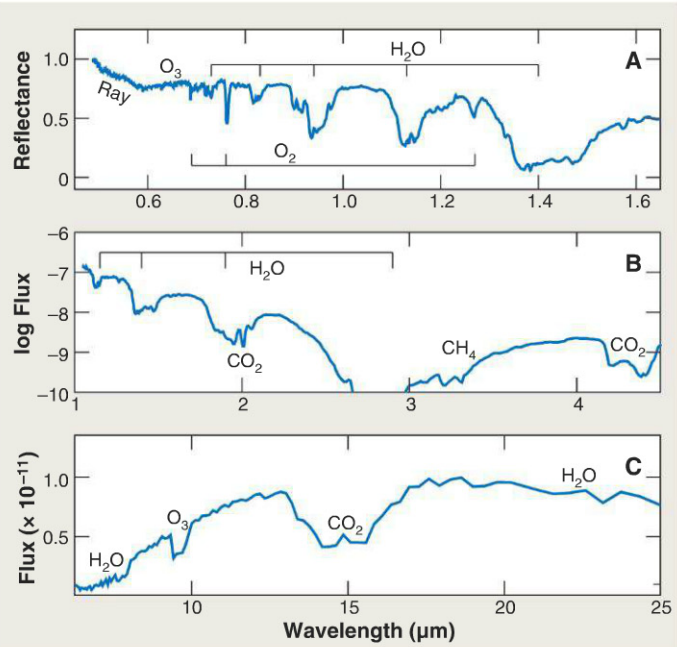


Fig. 3. Earth as an exoplanet, via observed disk-integrated spectra. (A) Visible-wavelength spectrum from Earthshine measurements plotted as normalized reflectance (*67*). **(B)** Near-infrared spectrum from NASA's EPOXI mission with flux in units of W m^{-2} μm^{-1} (*68*). **(C)** Mid-infrared spectrum as observed by Mars Global Surveyor enroute to Mars with flux in units of W m^{-2} Hz^{-1} (*69*). Major molecular absorption features are noted, including Rayleigh scattering. Only Earth's spectroscopically active, globally mixed gases would be observable from a remote space telescope.

Biosignature Gases

The main interest in defining a habitable planet is to identify an inhabited one, via remote-sensing observations of biosignature gases. Biosignature gases are gases produced by life that accumulate in a planetary atmosphere to high enough levels for remote detection by futuristic space telescopes. The underpinning assumption is that life uses chemical reactions to extract, store, and release energy, such that biosignature gases are generated as by-products somewhere in life's metabolic process.

Atmospheric biosignature gases have been studied theoretically as indicators of life for nearly half a century (*40*, *41*), with the proposed concept that a favorable biosignature gas is one that is many orders of magnitude out of thermochemical equilibrium with the planetary atmosphere.

Not all biosignature gases will be detectable from afar. Only globally mixed, spectroscopically active gases will be visible in an exoplanet spectrum. On Earth, the dominant global biosignature gases are O_2 (and its photolytic product O_3) produced by plants and photosynthetic bacteria, N_2O, and for early Earth possibly CH_4 (*42*) (Fig. 3).

The microbial world on Earth is incredibly diverse, and microorganisms produce a broad range of gases (*43*). Some of these gases, such as CO_2, are not unique to life as they occur naturally in the atmosphere. Other biosignature gases may be negligible on present-day Earth but accumulate to relevant levels in an environment substantially different from Earth's. Some examples that

have been studied for terrestrial-like atmospheres include organosulfur compounds, particularly methanethiol (CH_3SH, the sulfur analog of methanol) (*44*); CH_3Cl, a hydrogen halide (*45*); and sulfur biogenic gases on anoxic planets (*46*).

A major highlight from the last decade of biosignature gas research is the realization that low-EUV radiation environments, compared to solar radiation levels, lead to a much higher concentration of biosignature gases. This is because the stellar EUV radiation creates the radical OH (in some cases O), which destroys many gases in the atmosphere and thus reduces the gas lifetime (*45*). In an H_2-rich atmosphere, the same result holds with H as the major reactive species. Low-EUV radiation environments, compared to solar radiation levels, are found around inactive M dwarf stars (*47*).

Many biosignature gases have a "false positive" interpretation because they can be produced abiotically. False positives can, it is hoped, be identified by other atmospheric diagnostics. For example, photodissociation of water vapor in a runaway greenhouse with H escaping to space could lead to detectable O_2 levels. This situation could be identified by an atmosphere heavily saturated with water vapor. O_2 could also accumulate in a dry, CO_2-rich planet with weak geochemical sinks for O_2, a case that could be identified through strong CO_2 and weak H_2O features (*48*, *49*).

A sobering thought usually left unacknowledged is that when we finally discover biosignature gases, it may be not with a triumphant 100% certainty but rather with an assigned probability, depending on the level at which the false positive likelihood can be ascertained.

How to Find and Identify a Habitable World

In parallel to developing the theoretical foundation for planetary habitability, astronomers are developing instruments, telescopes, and space mission concepts to find and identify habitable or inhabited worlds. There are two ways to observe exoplanet atmospheres, and this leads to a "two-pronged approach."

The first approach is direct imaging [reviewed in (*50*)]. Here, the planet is observed as a point source (not spatially resolved like the beautiful Apollo images of Earth), and with the appropriate instrumentation, the light could be dispersed into a spectrum. The two objectives are to spatially separate the planet and star on the sky and to observe the planet literally within the glare of the host star. The limiting challenge for a planet like Earth is not its faintness—a relatively nearby Earth would not be fainter than the faintest galaxies ever observed by the Hubble Space Telescope—but the planet's proximity to a bright host. The Sun is 10 billion times as bright as Earth at visible wavelengths. The low-luminosity M stars are even more challenging to observe because of the smaller planet-star angular separation on the sky for the habitable zone. The use of a space-based telescope to image these planets is essential, both to get above the blurring effects of Earth's atmosphere and to avoid having to contend with the presence of these gases in our own atmosphere during an Earth-based hunt for biosignatures [compare (*51*)]. Implementation of the optical mathematics and engineering for blocking out starlight for planet finding is a subfield that has proceeded at a breathtaking pace (*50*), culminating in many concepts described under the umbrella term "Terrestrial Planet Finder" (TPF) (named after a cancelled set of missions under study by NASA in the early 2000s; the European Space Agency had a version called "Darwin"). Although a spectroscopically capable direct-imaging space mission to survey the 100 nearest Sun-like stars is now out of reach owing to an estimated cost of more than 5 billion dollars, technology development is still ongoing (*52*). A prescient saying in the exoplanet community Is that "all roads lead to TPF," because space-based direct imaging is the prime way to find and identify a true Earth twin.

The second approach is transit finding [reviewed in (*53*)] and transit spectroscopy. When a planet goes in front of its host star as seen from a telescope, some of the starlight will pass through the planet's atmosphere, and the atmospheric features will be imprinted on the starlight. In addition, when the planet goes behind the star (called "secondary eclipse"), the planet light will disappear and then reappear. For such transit and eclipse observations, the planet and star are not spatially separated on the sky but are instead observed in the "combined light" of the planet-star system: Using the starlight as a calibration tool enables the high-contrast measurements. Atmospheres of dozens of hot Jupiter exoplanets have been observed in this way. Although the Earth-Sun analog signal is still too small for observation, Earth-size and larger planets transiting M stars are suitable (*54*). M stars are favorable in many ways, from detectability to characterization, because the small star makes relative planet-to-star measurement signals larger than for Sun-like stars (*55*).

The obstacle to observing transiting planets is that the required orbital alignment will be fortuitous and infrequent, limiting the numbers of transiting planets accessible for study. The good news is that for planets orbiting quiet M stars, biosignature gases will accumulate, and simulations show that several such objects should exist and will be available for study with the under-construction James Webb Space Telescope [e.g., (*56*)]. First, we need a pool of suitable transiting planets orbiting quiet M stars (*57*) and next, a large amount of telescope time, perhaps tens of hours or more per planet. This scenario represents our nearest-term chance of identifying a habitable world.

Epilogue

Planet habitability is planet specific, even with the main imposed criterion that surface liquid water must be present. This is because the huge range of planet diversity in terms of masses, orbits, and star types should extend to planet atmospheres and interiors, based on the stochastic nature of planet formation and subsequent evolution. The diversity of planetary systems extends far beyond planets in our solar system. The habitable zone could exist from about 0.5 AU out to 10 AU for a solar-type star, or even beyond, depending on the planet's interior and atmosphere characteristics. As such, there is no universal habitable zone applicable to all exoplanets.

Many questions related to physical processes that govern the atmosphere, which itself controls habitability, may remain unanswered owing to a lack of observables. For example, which planets have plate tectonics and which have protective magnetic fields? Either there are no connections to observables or the observables are too weak for current and future instrumentation to measure.

Research strides are currently being made with statistical assessments of the occurrence rate of different sizes and masses of planets. This statistical phase of exoplanet research is moving toward estimates of the frequency of habitable planets with a handful of habitable-zone candidates tentatively identified. This statistical phase of exoplanets is expected to continue to flourish and dominate exoplanet science until the next generation of ground- and space-based telescopes.

Ultimately, a return to study of compelling individual objects is required—at any cost—if we want to assess a planet's habitability or attain the goal of identifying signs of life via biosignature gases. Is there any hope that the next space telescope, the James Webb Space Telescope, could be the first to provide evidence of biosignature gases? Yes, if—and only if—every single factor is in our favor. First, we need to discover a pool of super-Earths transiting in the "extended" habitable zones of nearby, quiet M stars. Second, life must not only exist on one of those planets, but must also produce biosignature gases that are spectroscopically active. Regardless of the search for life, the field of exoplanet characterization is on track to understand habitability and to find habitable worlds.

References and Notes

1. M. Mayor, D. Queloz, *Nature* **378**, 355 (1995).
2. A. Cassan *et al.*, *Nature* **481**, 167 (2012).
3. F. Fressin *et al.*, http://arxiv.org/abs/1301.0842 (2013).
4. A. W. Howard *et al.*, *Science* **340**, 572 (2013).
5. C. D. Dressing, D. Charbonneau, http://arxiv.org/abs/1302.1647 (2013).
6. S. S. Huang, *Am. Sci.* **47**, 397 (1959).
7. M. H. Hart, *Icarus* **33**, 23 (1978).
8. J. C. G. Walker, P. B. Hays, J. F. Kasting, *J. Geophys. Res.* **86**, 9776 (1981).
9. J. F. Kasting, D. P. Whitmire, R. T. Reynolds, *Icarus* **101**, 108 (1993).
10. R. K. Kopparapu *et al.*, http://arxiv.org/abs/1301.6674 (2013).
11. S. Lubow, S. Ida, in *Exoplanets*, S. Seager, Ed. (Univ. of Arizona Press, Tucson, AZ, 2011), p. 347.
12. J. Baross *et al.*, Committee on the Limits of Organic Life in Planetary Systems, Committee on the Origins and Evolution of Life, National Research Council (National Academies Press, Washington, DC, 2007).

13. W. Bains, *Astrobiology* **4**, 137 (2004).
14. The early Earth's habitability is not fully understood, owing to a phenomenon dubbed the "faint young sun paradox." A few billion years ago, the Sun was 20 to 30% fainter than it is today, based on asteroseismology-constrained stellar evolution models (*58*). Yet, there is no evidence that Earth was frozen over during that time, and the mechanism (including the possibility of a higher concentration of atmospheric greenhouse gases) for keeping Earth warm is not fully agreed upon.
15. C. de Bergh *et al.*, *Science* **251**, 547 (1991).
16. Some argue that Venus never had surface liquid water or the runaway greenhouse stage. One alternate explanation is that the noble gas isotopic ratios and low atmospheric O_2 are best explained by an evolutionary model where the dense CO_2 atmosphere forms very early on by magma crystallization (*59*).
17. C. I. Fassett, J. W. Head III, *Icarus* **198**, 37 (2008).
18. Models in (*10*) do not include water clouds. Water clouds generally have a cooling effect on planet atmospheres and would extend the habitable zone inward of <0.99 AU.
19. Liquid surface water may be detectable by the polarized ocean glint (specular reflection) with very futuristic space telescopes (*60*). In addition, variable visible-wavelength brightness attributed to clouds may indicate water oceans (*61*).
20. D. Deming *et al.*, http://adsabs.harvard.edu/abs/2013arXiv1302.1141D (2013).
21. J. L. Bean, E. M. Kempton, D. Homeier, *Nature* **468**, 669 (2010).
22. L. A. Rogers, S. Seager, *Astrophys. J.* **716**, 1208 (2010).
23. W. Wagner, A. Pruss, *J. Phys. Chem. Ref. Data* **31**, 387 (2002).
24. Lab-cultured microorganisms have been observed at temperatures as high as 395 K (*62*), and evidence for intact microorganism DNA and RNA has been seen at 473 K (*63*). A slightly higher value of up to 500 K may be considered for stability of biomolecules (*64*).
25. T. M. Seward, E. U. Franck, *Ber. Bunsenges. Phys. Chem* **85**, 2 (1981).
26. L. T. Elkins-Tanton, S. Seager, *Astrophys. J.* **685**, 1237 (2008).
27. H. Lammer *et al.*, *Earth Planets Space* **64**, 179 (2012).
28. C. O'Neill, A. Lenardic, *Geophys. Res. Lett.* **34**, 1 (2007).
29. D. Valencia, R. J. O'Connell, D. D. Sasselov, *Astrophys. J.* **670**, 45 (2007).
30. J. F. Kasting, D. Catling, *Annu. Rev. Astron. Astrophys.* **41**, 429 (2003).
31. L. Schaefer, B. Fegley Jr., *Icarus* **208**, 438 (2010).
32. A. Borysow, *Astron. Astrophys.* **390**, 779 (2002).
33. A. Borysow, L. Frommhold, G. Birnbaum, *Astrophys. J.* **326**, 509 (1988).
34. R. Pierrehumbert, E. Gaidos, *Astrophys. J.* **734**, L13 (2011).
35. D. J. Stevenson, *Nature* **400**, 32 (1999).
36. Y. Abe, A. Abe-Ouchi, N. H. Sleep, K. J. Zahnle, *Astrobiology* **11**, 443 (2011).
37. A. Zsom, S. Seager, J. de Wit, http://arxiv.org/abs/1304.3714 (2013).
38. S. Raymond, T. Quinn, J. I. Lunine, *Icarus* **168**, 1 (2004).
39. A relevant note is that scorching Mercury has water ice at its poles (*65*).
40. J. Lederberg, *Nature* **207**, 9 (1965).
41. J. E. Lovelock, *Nature* **207**, 568 (1965).
42. D. J. Des Marais *et al.*, *Astrobiology* **2**, 153 (2002).
43. S. Seager, M. Schrenk, W. Bains, *Astrobiology* **12**, 61 (2012).
44. C. B. Pilcher, *Astrobiology* **3**, 471 (2003).
45. A. Segura *et al.*, *Astrobiology* **5**, 706 (2005).
46. S. D. Domagal-Goldman, V. S. Meadows, M. W. Claire, J. F. Kasting, *Astrobiology* **11**, 419 (2011).
47. At 100 to 200 nm, a quiet M star EUV flux is about a factor of 1000 lower than a solar-like star's flux. At 200 to 300 nm, the quiet M star EUV flux is about a factor of 100 lower.
48. F. Selsis, D. Despois, J.-P. Parisot, *Astron. Astrophys.* **388**, 985 (2002).
49. A. Segura, V. S. Meadows, J. F. Kasting, D. Crisp, M. Cohen, *Astron. Astrophys.* **472**, 665 (2007).
50. W. A. Traub, B. R. Oppenheimer, in *Exoplanets*, S. Seager, Ed. (Univ. of Arizona Press, Tucson, AZ, 2011), p. 111.
51. I. A. G. Snellen, R. J. de Kok, R. le Poole, M. Brogi, J. Birkby, *Astrophys. J.* **764**, 182 (2013).
52. http://exep.jpl.nasa.gov/technology/
53. J. Winn, in *Exoplanets*, S. Seager, Ed. (Univ. of Arizona Press, Tucson, AX, 2011), p. 55.
54. L. Kaltenegger, W. A. Traub, *Astrophys. J.* **698**, 519 (2009).
55. P. Nutzman, D. Charbonneau, *Publ. Astron. Soc. Pac.* **120**, 317 (2008).
56. D. Deming *et al.*, *Publ. Astron. Soc. Pac.* **121**, 952 (2009).
57. NASA recently selected the Transiting Exoplanet Survey Satellite (TESS) for launch in 2017. TESS is an all-sky survey that will discover thousands of exoplanets orbiting nearby stars, including the highly prized habitable-zone planets transiting M stars, the pool of planets suitable for atmospheric follow-up with the James Webb Space Telescope.
58. J. N. Bahcall, M. H. Pinsonneault, S. Basu, *Astrophys. J.* **555**, 990 (2001).
59. C. Gillmann, E. Chassefière, P. Lognonné, *Earth Planet. Sci. Lett.* **286**, 503 (2009).
60. D. M. Williams, E. Gaidos, *Icarus* **195**, 927 (2008).
61. E. B. Ford, S. Seager, E. L. Turner, *Nature* **412**, 885 (2001).
62. K. Takai *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 10949 (2008).
63. M. O. Schrenk, D. S. Kelley, J. R. Delaney, J. A. Baross, *Appl. Environ. Microbiol.* **69**, 3580 (2003).
64. E. W. Lang, *Adv. Space Res.* **6**, 251 (1986).
65. D. J. Lawrence *et al.*, *Science* **339**, 292 (2013).
66. H. Rein, http://arxiv.org/abs/1211.7121 (2012).
67. M. C. Turnbull *et al.*, *Astrophys. J.* **644**, 551 (2006).
68. T. D. Robinson *et al.*, *Astrobiology* **11**, 393 (2011).
69. P. R. Christensen, J. C. Pearl, *J. Geophys. Res.* **102**, 10875 (1997).

Acknowledgments: I thank A. Zsom, V. Stamenkovic, and L. Rogers for useful discussions and D. Beckner for producing the figures.

10.1126/science.1232226

LIVE CELL IMAGING:

The Future for Discoveries

WEBINAR

REGISTER NOW!
webinar.sciencemag.org

Cellular imaging remains one of the most important techniques in life science research based in part on the notion that "seeing is believing." For a variety of practical reasons including types of probes available and access of probes to their intracellular targets, imaging is currently most often performed on specimens that have been fixed and labeled. Although technically more challenging, live cell imaging enables researchers to study cellular events and processes that cannot be visualized in fixed specimens and to follow short- and long-term dynamic processes to gain deeper insights into the complex mechanisms of cell biology.

During this webinar, viewers will:

- Learn from our speakers the benefits of imaging live cells using techniques such as high resolution microscopy, superresolution microscopy, and high-content analysis
- Receive practical advice on how to overcome some of the technical challenges in live cell imaging
- Obtain guidance on how they can modify their imaging practices to obtain superior results
- Have their questions answered by our respected thought leaders.

Wednesday, May 22, 2013
12 noon Eastern, 9 a.m. Pacific,
5 p.m. UK, 6 p.m. Central Europe

SPEAKERS:

Lynne Turnbull, Ph.D.
University of Technology
Sydney, Australia

Edward M. Campbell, Ph.D.
Loyola University
Chicago, IL

Nick Thomas, Ph.D.
GE Healthcare
Cardiff, Wales

Sponsored by
GE Healthcare

Brought to you by the
Science/AAAS Custom
Publishing Office

Emergence and Frustration of Magnetism with Variable-Range Interactions in a Quantum Simulator

R. Islam,[1]* C. Senko,[1] W. C. Campbell,[1]† S. Korenblit,[1] J. Smith,[1] A. Lee,[1] E. E. Edwards,[1] C.-C. J. Wang,[2]‡ J. K. Freericks,[2] C. Monroe[1]§

Frustration, or the competition between interacting components of a network, is often responsible for the emergent complexity of many-body systems. For instance, frustrated magnetism is a hallmark of poorly understood systems such as quantum spin liquids, spin glasses, and spin ices, whose ground states can be massively degenerate and carry high degrees of quantum entanglement. Here, we engineer frustrated antiferromagnetic interactions between spins stored in a crystal of up to 16 trapped $^{171}Yb^+$ atoms. We control the amount of frustration by continuously tuning the range of interaction and directly measure spin correlation functions and their coherent dynamics. This prototypical quantum simulation points the way toward a new probe of frustrated quantum magnetism and perhaps the design of new quantum materials.

Predicting the behavior of many-body quantum materials such as frustrated magnets is difficult because the number of relevant configurations scales exponentially with the number of particles (*1*–*3*). Feynman proposed the use of a quantum simulator for the task. Here, interactions are engineered in a "standard" quantum system to illuminate the physics behind the real material (*4*). Cold atoms are excellent standards for quantum simulations of magnetism, with the ability to tailor frustrated magnetic interactions and measure the individual atomic spins (*5*, *6*). Neutral atomic systems are typically limited to nearest-neighbor interactions (*7*), although geometrically frustrated interactions can be realized in certain optical lattice geometries (*8*). The strong Coulomb interaction between cold atomic ions (*9*) has led to the realization of long-range Ising couplings between individual trapped ion spins (*10*–*14*) and the observation of spin frustration and quantum entanglement in the smallest system of three spins (*15*).

Here, we report the implementation of variable-range antiferromagnetic (AFM) Ising interactions with up to 16 atomic ion spins, using optical dipole forces. We directly measure the emergence and frustration of magnetic ordering through spatially resolved imaging of the ions. The spins are initially polarized along a strong effective magnetic field transverse to the Ising couplings, and when the field is lowered, the spins order themselves according to the characterstics of the Ising interactions. We can increase the level of frustration by increasing the range of the interactions, which results in a more equitable balance of competing interactions and a suppression of magnetic order. The quantum coherence in the system is characterized by reversing the transverse field back to its initial value and comparing the resulting state with the initial state. These experiments present simulations in a nontrivial quantum system that approaches a complexity level at which it becomes difficult or impossible to calculate the spin dynamics.

[1]Joint Quantum Institute, University of Maryland Department of Physics and National Institute of Standards and Technology, College Park, MD 20742, USA. [2]Department of Physics, Georgetown University, Washington, DC 20057, USA.

*Present address: Department of Physics, Harvard University, Cambridge, MA 02138, USA.
†Present address: Department of Physics and Astronomy, University of California, Los Angeles, CA 90095, USA.
‡Present address: Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA.
§Corresponding author. E-mail: monroe@umd.edu

Model

We simulate the quantum transverse Ising model with long-range AFM interactions in a one-dimensional spin chain, given by the Hamiltonian

$$H = \sum_{j<i} J_{ij}\sigma_x^{(i)}\sigma_x^{(j)} - B\sum_i \sigma_y^{(i)} \qquad (1)$$

where the Planck constant h is set to 1, $\sigma_\gamma^{(i)}$ ($\gamma = x, y, z$) are the spin-1/2 Pauli operators for the ith spin ($i = 1,2,\ldots N$), B is the effective transverse magnetic field, and $J_{ij} > 0$ is the Ising coupling between spins i and j, falling off with the lattice spacing $|i - j|$ approximately as

$$J_{ij} \approx \frac{J_0}{|i-j|^\alpha}, \qquad (2)$$

where $0 < \alpha < 3$ (*8*).

For $B \gg J_{ij}$ on all pairs, the spins are polarized along the effective transverse magnetic field in the ground state $|\uparrow_y \uparrow_y \uparrow_y \ldots\rangle$ of the Hamiltonian in Eq. 1, where $|\uparrow_y\rangle$ denotes a spin along the $+y$ direction of the Bloch sphere. As the ratio of B to the Ising couplings is reduced, the system crosses over to an ordered state dictated by the form of the Ising couplings, and the spectrum of energy levels depends on the range of the interactions. For any finite $\alpha > 0$, we find by direct calculation that the staggered AFM states $|\uparrow\downarrow\uparrow\downarrow \ldots\rangle$ and $|\downarrow\uparrow\downarrow\uparrow \ldots\rangle$ constitute the doubly degenerate ground state manifold at $B = 0$, with the degeneracy arising from the time reversal or the global spin flip symmetry of the Hamiltonian. Here $|\uparrow\rangle$ and $|\downarrow\rangle$ are spin states oriented along the Ising or x direction of the Bloch sphere. Thus, the system exhibits nearest-neighbor AFM or Néel ordering at sufficiently low temperatures. When the interactions are uniform over all pairs of spins ($\alpha \to 0$), the system becomes maximally frustrated

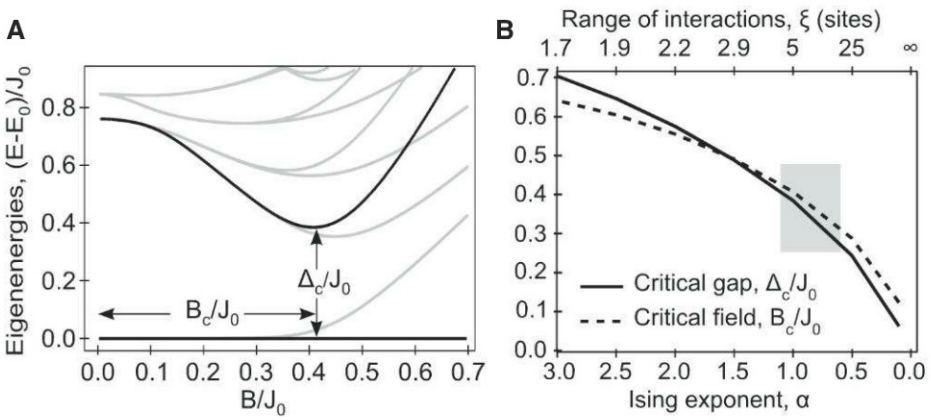


Fig. 1. Theoretical energy spectrum and critical gap in the long-range antiferromagnetic Ising model (Eq. 1) for $N = 10$ spins. (A) Low-lying energy states for $\alpha = 1$ (characteristic range of $\xi = 5$ sites) as a function of the dimensionless parameter B/J_0. The spacing between the ground state at $E = E_0$ and the first coupled excited state (black lines) reaches a bottleneck at a critical value B_c/J_0 with critical gap Δ_c. **(B)** Theoretical dependence of B_c/J_0 (dotted line) and Δ_c/J_0 (solid line) on the range of the interaction. As the interaction range increases, the competing long-range couplings make it easier to create excitations and the critical gap is reduced, so a relatively small effective transverse field can break the spin ordering. Both parameters approach zero as $\alpha \to 0$ or $\xi \to \infty$. Present experiments are performed with parameters in the shaded region.

and the excitation gap (Fig. 1A) closes, leading to a finite entropy density in the ground state. In this case, any spin configuration with a net magnetization of zero (1/2) for even (odd) numbers of spins belongs to the ground state.

Between the limits $B = 0$ and $B >> J_{ij}$, the energy spectrum features a minimum gap, whose position and size depends on the degree of frustration in the system, or the interaction range. (The interaction range is defined as the number of lattice spacings ξ where the interaction falls off to 20% of the nearest-neighbor Ising coupling: $\xi \equiv 5^{1/\alpha}$.) Figure 1A shows a few low-lying energy states of the Hamiltonian in Eq. 1 for an interaction range of $\xi = 5$ (corresponding to $\alpha = 1$). The first excited eigenstate merges with the ground state for small B/J_0 and has the same spin order as the ground state near $B/J_0 = 0$. The critical gap Δ_c between the ground and the first coupled excited state determines the adiabaticity criterion (*16*). Figure 1B compares the position (dotted line) and size (solid line) of the critical gap of the Hamiltonian for various ranges. As the range and hence the amount of frustration increases, the critical field is pushed toward zero, and the gap closes. To observe the effects of frustration, reflected in the density of states near the ground state, we quench the system by ramping the effective transverse magnetic field faster than the critical gap ($|\dot{B}/B| > \Delta_c$) to populate the lowest coupled excited states. The observed spin order depends on the resulting degree of excitation and hence on the level of frustration.

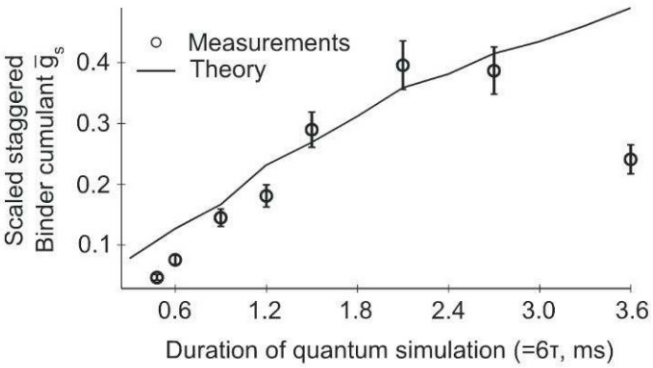


Fig. 2. Spin order versus speed of ramp, for $N = 10$ spins. The spins are initialized with $B/J_0 = 5$ and the transverse field is ramped exponentially down for six time constants, and the experiment is repeated for various values of time constant τ. Symbols (solid line) indicate the scaled staggered Binder cumulant $\bar{g}_s$, versus the total duration 6τ of the ramp measured (expected from theory). As the Hamiltonian evolves more slowly, the observed spin order shows more ground state order, and less excitation for ramp times under ~ 2.5 ms. For longer times, the spins become disordered, implying external decoherence in the system. Here (and in the following figures) the error bars include statistical fluctuations and estimated detection uncertainties.

Experiment

The spins are realized in a collection of $^{171}Yb^+$ ions confined in a linear radiofrequency (Paul) trap, with the effective spin-1/2 system represented by two hyperfine "clock" states within each ion $|\uparrow_z\rangle$ and $|\downarrow_z\rangle$, separated by the hyperfine frequency $\nu_{HF} = 12.642819$ GHz (*17*). The variable-range AFM Ising interactions are generated by applying off-resonant spin-dependent optical dipole forces (*11*) that drive stimulated Raman transitions between the spin states while modulating the Coulomb interaction between the ions in a controlled way (*18*). The effective magnetic field is generated by simultaneously driving coherent transitions between the spin states with a $\pi/2$-phase shift with respect to the dipole force beams. At any time, we measure the state of the spins by illuminating the ions with resonant radiation and collecting state-dependent fluorescence on an imager with site-resolving optics (*17*). From this information, we can extract all spin correlation functions (*18*).

The quantum simulation begins by optically pumping all spins to the $|\downarrow_z\rangle$ state and then coherently rotating them all about the x axis of the Bloch sphere to initialize each spin in state $|\uparrow_y\rangle$ along the effective transverse magnetic field.

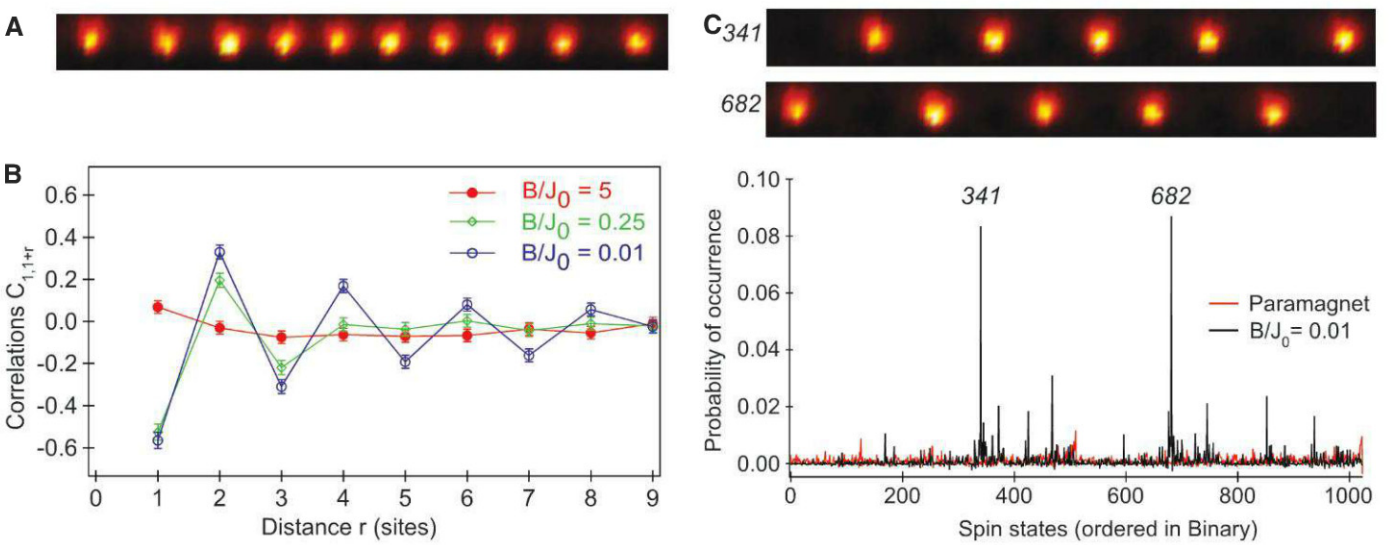


Fig. 3. Quantum phase transition from a paramagnet to an antiferromagnet (Eq. 1), with $J_{ij} \approx |i - j|^{-1.05}$ in a system of 10 spins. The exponent ($\alpha = 1.05$) is estimated from the average couplings between spins in this inhomogeneous system. (**A**) Image of 10 trapped ions, with a distance of 22 μm between the first and last ion. (**B**) Measured two-point correlation function between a chosen spin (on the edge) and the other spins separated by r lattice sites, $C_{1,1+r} = \langle\sigma_x^{(1)}\sigma_x^{(1+r)}\rangle - \langle\sigma_x^{(1)}\sigma_x^{(1+r)}\rangle$, averaged over 4000 experiments for each value of the parameter B/J_0. For $B/J_0 = 5$, the spins are initially polarized along the transverse y field with little correlation along the Ising x direction. As the field is reduced, the spins cross over to predominantly AFM states $|\uparrow\downarrow\uparrow\downarrow\ldots\rangle$ and $|\downarrow\uparrow\downarrow\uparrow\ldots\rangle$, resulting in alternating signs in the two-point correlations with separation. The solid lines are shown to guide the eye. (**C**) Measured occurrence probability of all $2^{10} = 1024$ states at $B/J_0 \approx 5$ (paramagnetic state, red trace) and $B/J_0 \approx 0.01$ (AFM phase, black trace). The states are listed in binary order, with spin $|\downarrow\rangle \equiv 0$ and $|\uparrow\rangle \equiv 1$. The residual peaks in the red trace are consistent with detection errors biased toward states with many $|\uparrow\rangle$ spins such as 127, 511, and 1023. The two tall peaks in the black trace at 341 (0101010101) and 682 (1010101010) correspond to two Néel-ordered staggered AFM states, shown with camera images of these cases and contributing ~17% to the population.

The Hamiltonian (Eq. 1) is then switched on with an initial field $B_0 \approx 5\ J_0$, where J_0 (~1 kHz) is the average nearest-neighbor Ising coupling, thus preparing the spins in the ground state of the initial Hamiltonian with a fidelity better than 97%. The effective magnetic field is ramped down exponentially in time with a time constant of 400 μs to a final value B of the transverse field, but no longer than 2.4 ms overall, to avoid decoherence effects present in the system (see Fig. 2). At this point, the Hamiltonian is switched off, freezing the spins for measurement. We then measure the x or y component of each spin $\sigma_x^{(i)}$ or $\sigma_y^{(i)}$ by first rotating our measurement axes with an appropriate global π/2 pulse similar to the initialization procedure, before capturing the spin-dependent fluorescence on the imager. The experiments are repeated 2000 to 4000 times to measure expectation values of certain spin operators and correlation functions (denoted by ⟨... ⟩).

Order Parameters and Correlation Functions

From these measurements, we can construct order parameters appropriate for observing low-energy AFM states. Various moments of the staggered magnetization operator $m_s = \frac{1}{N}\sum_{i=1}^{N}(-1)^i\sigma_x^{(i)}$ distinguish between paramagnetic and AFM order, and also quantify spin flip excitations. In particular, the normalized fourth moment of this magnetization, known as the Binder cumulant (*19*) $g_s = \langle(m_s - \langle m_s\rangle)^4\rangle/\langle(m_s - \langle m_s\rangle)^2\rangle^2$, scaled to $\overline{g}_s = (g_s^0 - g_s)/(g_s^0 - 1)$ to remove finite size effects, varies from $\overline{g}_s = 0$ to $\overline{g}_s = 1$ as the paramagnetic state gives way to AFM order (see Fig. 2). Here $g_s^0 = 3 - 2/N$ is the Binder cumulant in a perfect paramagnetic state of the N spins. We can also form any correlation function of the spins such as the two-point correlation $C_{i,j} = \langle\sigma_x^{(i)}\sigma_x^{(j)}\rangle - \langle\sigma_x^{(i)}\rangle\langle\sigma_x^{(j)}\rangle$, allowing a direct probe of spin order for each experimental realization. The Fourier transform of this correlation function is the structure function $S(k) = \frac{1}{N-1}\left|\sum_{r=1}^{N-1}C(r)e^{ikr}\right|$, where $C(r) = \frac{1}{N-r}\sum_{m=1}^{N-r}C_{m,m+r}$ is the average correlation over r sites in the chain. The structure function shows spin order versus wave number k, with $S(k = \pi)$ singling out the presence of the nearest-neighbor Néel AFM order.

Figure 3 shows the onset of AFM ordering in the quantum simulation of the frustrated transverse field Ising model in a system of 10 spins. Two-point spin correlations $C_{1,r}$ between a chosen edge spin and the other spins are presented in Fig. 3B at various stages in the ramp $B/J_0 = 5$, 0.25, and 0.01. For larger transverse magnetic fields, there are no appreciable correlations between the spin components along the Ising direction. As the ratio of B/J_0 is lowered, however, a zig-zag pattern emerges, with negative (positive) correlations between spins separated by odd (even) lattice spacings. For $B/J_0 \approx 0.01$, the nearest-neighbor spin correlation reaches about 60%, and the correlation length (defined to be the distance at which the absolute correlation drops to $1/e$ of the nearest neighbor value) reaches about six lattice sites. The effective field was ramped exponentially down from $B/J_0 \approx 5$ with a time constant of 400 μs in this experiment. This ramping is not slow enough to be adiabatic, and the diabatic effects prevent the spin ordering from reaching a perfect AFM phase. Figure 3C shows the measured probabilities of all $2^{10} = 1024$ possible spin states measured along the Ising x direction, sorted in binary order with spin $|\downarrow\rangle \equiv 0$ and $|\uparrow\rangle \equiv 1$. The net detection fidelity of each spin is ~97%, after postfiltering the measurements based on calibrating the known detection errors for each spin (*18*, *20*). The initial paramagnetic phase shown in red ($B/J_0 \approx 5$) exhibits a roughly uniform probability of $1/1024 \approx 0.1\%$ for each state (the residual peaks in the red trace are consistent with detection errors). The spin-ordered phase shown in black ($B/J_0 = 0.01$) displays the emergence of the two AFM states, each with an occupation probability of about 8.5%. Other prominent peaks correspond to single spin-flips and other low-lying excitations from the two ground states.

In Fig. 4, we probe the frustration in the system for various ranges of interactions. Here, we look at the spin order achieved in the quantum simulation when the external magnetic field is ramped down to $B/J_0 \approx 0.01$ for four different ranges of interactions. In Fig. 4A, we show the measured structure function $S(k)$ at wave vector values $k = \frac{\pi}{10}, \frac{2\pi}{10}, \ldots, \pi$ from the measured two-point correlation functions. To directly compare the different interaction ranges, we choose the same external magnetic field ramp time constant, $\tau = 0.4/J_0$. As the range of interaction increases, the ground state AFM order (given by the structure function peak at $k = \pi$) disappears, reflecting increased occupation of the excited states as the frustration grows. Figure 4B displays the observed distribution of energy $P(E_i)$ for $\alpha = 1.05$ (shorter range) and $\alpha = 0.76$ (longer range) power-law exponents, along with the cumulative energy distribution function. The eigenenergies of each configuration are calculated using Eq. 1 with $B = 0$. For the longer-range interactions, excitations are more prevalent, and the energy gap between the ground and the first excited state is reduced, both of which are signatures of increasing frustration in the system. The observed final entropy per particle $S = -\frac{1}{N}\sum_i P(E_i)\log P(E_i)$ is seen to increase from 0.832 to 0.903 as the interaction range grows from $\alpha = 1.05$ ($\xi = 4.6$ sites) to $\alpha = 0.67$ ($\xi = 11$ sites), which is also a signature of the increased frustration in the system. As a reference, the paramagnetic state distribution shows an entropy per particle of 0.959, which is slightly less than unity because of detection errors.

Quantum Coherence

The above measurements of the state distribution concern only the diagonal components or

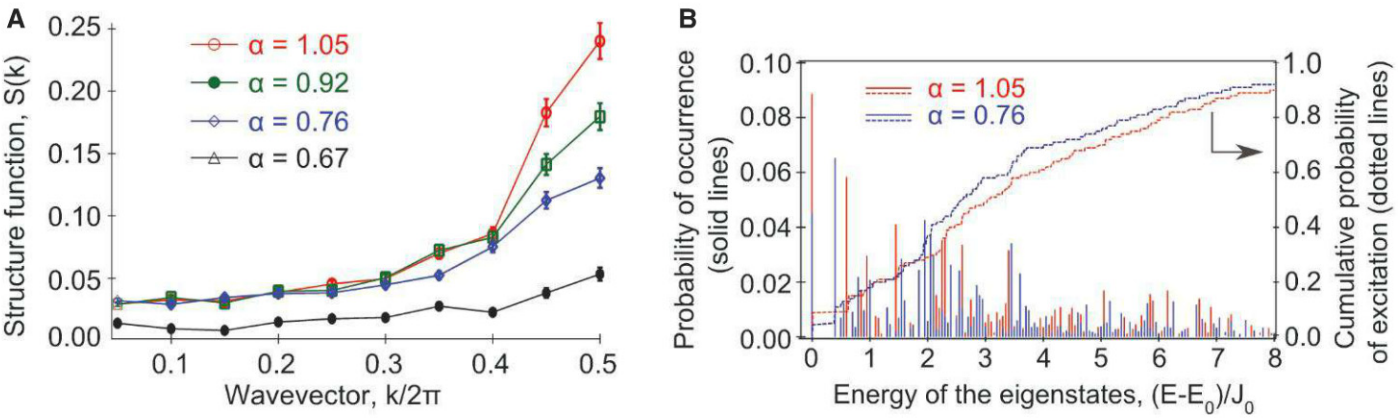


Fig. 4. (A) Structure function $S(k)$ versus wave vector k for various ranges of AFM interactions, for $B/J_0 = 0.01$ in a system of $N = 10$ spins. The increased level of frustration for the longer-range interactions reduces the observed antiferromagnetic spin order. The detection errors may be larger than shown here for the longest range of interactions, owing to spatial crosstalk from their closer spacing. (**B**) Distribution of observed states in the spin system, sorted according to their energy E_i that was previously calculated by diagonalizing Eq. 1 with $B = 0$. Data are presented for two ranges (red for $\alpha = 1.05$ and blue for $\alpha = 0.76$). The dashed lines indicate the cumulative energy distribution functions for these two ranges.

populations of the density matrix. To characterize the quantum coherence in the simulation, we retrace the external magnetic field back to its initial value after ramping it down to almost zero and measure each spin along the transverse (y) magnetic field. Figure 5 shows the distribution of the measured value m_y of the total transverse spin operator $S_y = \sum_i \sigma_y^{(i)}$ at three different times: first at the beginning of the simulation, second when the transverse field has been ramped down to nearly zero, and third after the transverse field returns to its initial value ($\alpha = 0.92$ for these data). The initial state is ideally a delta function at $\langle S_y \rangle = 10$, but finite detection efficiency broadens the distribution. At the lowest value of the field ($B/J_0 \approx 0.01$), the transverse magnetization is distributed near $\langle S_y \rangle = 0$, as the spins are presumably ordered along the Ising x direction. When the external field is ramped back to its initial state, the distribution of the total spin returns toward the initial distribution, with an average magnetization that is approximately $\langle S_y \rangle = 68(4)\%$ of its initial value, indicating that some level of quantum coherence is maintained throughout the simulation. This number is in agreement with the theoretically estimated average magnetization of $\langle S_y \rangle \approx 70\%$ of its initial value, obtained by numerically integrating the Schrödinger's equation without any decoherence.

To probe decoherence in the simulation, we repeat the experiment with various ramping speeds of the effective magnetic field. In Fig. 2, we plot the AFM order parameter g_s versus the total duration for the experiment for a long-range coupling ($\alpha = 1.05$) for $N = 10$ spins. Each data point represents the spin order achieved after ramping the magnetic field B down exponentially from $5J_0$ for a total duration of six time constants. The AFM order grows with slower ramping, as expected, for up to $\tau = 400\ \mu s$. But we also observe a saturation and then decay in the spin order, which might indicate the presence of decoherence in the system at long times. During the simulation, spontaneous Raman scattering from the optical beams is expected to occur at a rate of less than 6 s^{-1} per spin (*21*), which is consistent with separate measurements of the spin relaxation from a single spin and is therefore not expected to contribute to decoherence given the time scales in the experiment. The phonon population is expected to be well under 10% for all the data presented here (*22*). A principal source of decoherence appears to be the intensity fluctuations in the Raman beams, from beam pointing instabilities and fluctuations in the optical power.

Maintaining adiabaticity while ramping the magnetic field down is more difficult when the Ising interactions in the experiment are frustrated, because the relevant energy gaps are smaller. To directly compare frustrated versus nonfrustrated systems, we execute quantum simulations of both long-range AFM and ferromagnetic (FM) Ising models in a system of $N = 16$ spins (Fig. 6). For the FM experiment (Fig. 6C), we initialize the spins in the highest-energy state with respect to the transverse field $|\downarrow_y \downarrow_y \downarrow_y \ldots\rangle$ and ramp the field down as before (*10*, *15*). For the AFM experiment with the same ramp rate, we find that the best AFM nearest-neighbor correlation (Fig. 6A) is only 30(3)%, corresponding to a staggered magnetization of about 30(2)%, whereas the simulation of the FM model shows a clear FM spin order across the chain (Fig. 6B), reaching 73(3)% magnetization. We have also observed a level of 70(10)% FM magnetization emerging in $N = 18$ spins (*18*).

The interacting spin system that we study is approaching a complexity level at which it becomes difficult or impossible to predict its behavior. Static properties such as the ground state order or the excited state energies can be calculated for hundreds of spins using Monte Carlo methods (*23*); however, the calculation of dynamics of fully connected spin models is currently limited to approximately $N = 30$ spins (*24*, *25*).

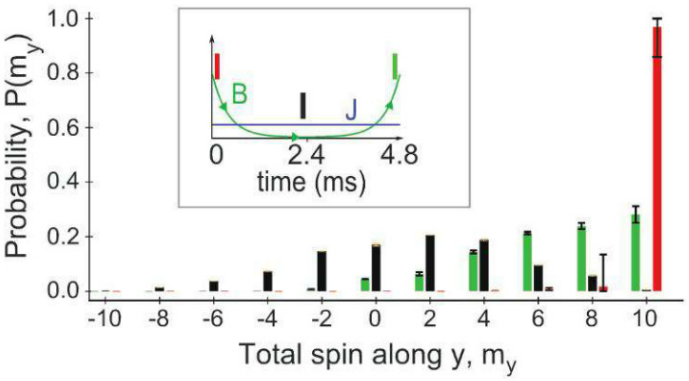


Fig. 5. Quantum coherence in the simulation, in a system of $N = 10$ spins. Probabilities of different values of the total spin component along y in the initial polarized state (red), when the transverse field is ramped to $\approx 0.01\ J_0$ (black), and when the transverse field is reversed back to its initial value (green). After reversal, the y magnetization returns to ~68(4)% of its initial value, indicating quantum coherence in the evolution. The trajectory of the transverse field (B, in green) and all the Ising couplings (J, in blue) are shown in the inset. $\alpha = 0.92$ for these data.

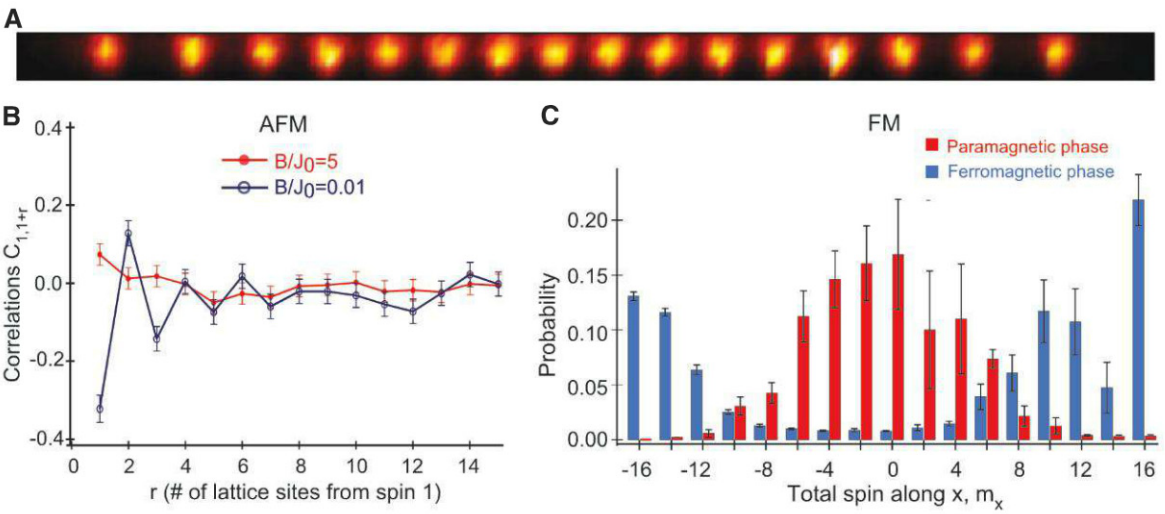


Fig. 6. Magnetic ordering in N=16 spins. (**A**) Image of 16 trapped ions, with a distance of 30 μm between the first and last ion. (**B**) Pair correlation function measured at various stages of the quantum simulation, for $B/J_0 = 5$ (red) and $B/J_0 = 0.01$ (blue) in an AFM coupling falling off with distance as $J_{ij} \sim |i - j|^{-1}$ among $N = 16$ spins. Small amounts of staggered order are seen, tempered by the small gaps and frustration in the low-energy states. (**C**) In contrast, for all FM couplings (again with $J_{ij} \sim |i - j|^{-1}$), the gaps are large and clear FM order is seen. Here the measured distribution of magnetization is plotted. The paramagnetic phase of 16 spins is indicated in red, and after the field is ramped to nearly zero, the distribution clearly bifurcates, indicating population weighted heavily toward the FM states $|\downarrow\downarrow\downarrow\ldots\rangle$ and $|\uparrow\uparrow\uparrow\ldots\rangle$, indicated in blue. The resulting magnitude magnetization is ~73%.

Small ion trap quantum simulators such as that reported here may soon reach this milestone with technical upgrades in the hardware, including lower vacuum chamber pressures to prevent collisions with the background gas, better stability of the optical intensities, and higher optical power so that fluctuations in the beam inhomogeneities can be suppressed.

References and Notes

1. H. T. Diep, *Frustrated Spin Systems* (World Scientific, Singapore, 2005).
2. R. Moessner, A. P. Ramirez, *Phys. Today* **59**, 24 (2006).
3. S. Sachdev, *Quantum Phase Transitions* (Cambridge Univ. Press, Cambridge, 1999).
4. R. Feynman, *Int. J. Theor. Phys.* **21**, 467 (1982).
5. I. Bloch, J. Dalibard, S. Nascimbéne, *Nat. Phys.* **8**, 267 (2012).
6. R. Blatt, C. F. Roos, *Nat. Phys.* **8**, 277 (2012).
7. J. Simon *et al.*, *Nature* **472**, 307 (2011).
8. G.-B. Jo *et al.*, *Phys. Rev. Lett.* **108**, 045305 (2012).
9. D. Porras, J. I. Cirac, *Phys. Rev. Lett.* **92**, 207901 (2004).
10. A. Friedenauer, H. Schmitz, J. T. Glueckert, D. Porras, T. Schaetz, *Nat. Phys.* **4**, 757 (2008).
11. K. Kim *et al.*, *Phys. Rev. Lett.* **103**, 120502 (2009).
12. R. Islam *et al.*, *Nat. Commun.* **2**, 377 (2011).
13. J. W. Britton *et al.*, *Nature* **484**, 489 (2012).
14. A. Khromova *et al.*, *Phys. Rev. Lett.* **108**, 220502 (2012).
15. K. Kim *et al.*, *Nature* **465**, 590 (2010).
16. E. E. Edwards *et al.*, *Phys. Rev. B* **82**, 060412 (2010).
17. S. Olmschenk *et al.*, *Phys. Rev. A* **76**, 052314 (2007).
18. See supplementary materials on *Science* Online.
19. K. Binder, *Phys. Rev. Lett.* **47**, 693 (1981).
20. C. Shen, L.-M. Duan, *New J. Phys.* **14**, 053053 (2012).
21. W. C. Campbell *et al.*, *Phys. Rev. Lett.* **105**, 090502 (2010).
22. C.-C. J. Wang, J. K. Freericks, *Phys. Rev. A* **86**, 032329 (2012).
23. R. Moessner, S. L. Sondhi, *Phys. Rev. B* **63**, 224401 (2001).
24. A. W. Sandvik, *Phys. Rev. Lett.* **104**, 137204 (2010).
25. Quantum Monte Carlo algorithms can be used to calculate static equilibrium properties of the transverse field Ising model for large numbers of interacting spins, so ground states and static correlation functions can indeed be calculated for large systems (*23*). However, the calculation of dynamics and nonequilibrium behavior of quantum spin models is not currently feasible for these Monte Carlo approaches, and in the presence of frustrated long-range interactions, the general behavior of such systems requires exact diagonalization. Other techniques such as the density matrix renormalization group become too difficult with long-range interactions. Because the size of the Hilbert space grows exponentially with the number of spins, sparse-matrix techniques such as the Lanczos method must therefore be used. Current state-of-the-art work on such systems is limited to sizes on the order of 30 to 35 spins (*24*).

Acknowledgments: We thank E. Demler, L. Duan, D. Huse, K. Kim, P. Richerme, R. Sensarma, and P. Zoller for critical discussions. This work is supported by the U.S. Army Research Office (ARO) award W911NF0710576 with funds from the Defense Advanced Research Projects Agency Optical Lattice Emulator Program, ARO award W911NF0410234 with funds from the Intelligence Advanced Research Projects Activity, and the NSF Physics Frontier Center at the Joint Quantum Institute. J.K.F. was also supported by the McDevitt bequest at Georgetown.

Supplementary Materials

www.sciencemag.org/cgi/content/full/340/6132/583/DC1
Supplementary Text
Fig. S1
References (*26–28*)

2 November 2012; accepted 8 February 2013
10.1126/science.1232296

REPORTS

Kepler-62: A Five-Planet System with Planets of 1.4 and 1.6 Earth Radii in the Habitable Zone

William J. Borucki,[1]* Eric Agol,[2] Francois Fressin,[3] Lisa Kaltenegger,[3,4] Jason Rowe,[5] Howard Isaacson,[6] Debra Fischer,[7] Natalie Batalha,[1] Jack J. Lissauer,[1] Geoffrey W. Marcy,[6] Daniel Fabrycky,[8,9] Jean-Michel Désert,[3] Stephen T. Bryson,[1] Thomas Barclay,[10] Fabienne Bastien,[11] Alan Boss,[12] Erik Brugamyer,[13] Lars A. Buchhave,[14,15] Chris Burke,[5] Douglas A. Caldwell,[5] Josh Carter,[3] David Charbonneau,[3] Justin R. Crepp,[16,17] Jørgen Christensen-Dalsgaard,[18] Jessie L. Christiansen,[5] David Ciardi,[19] William D. Cochran,[13] Edna DeVore,[5] Laurance Doyle,[5] Andrea K. Dupree,[3] Michael Endl,[13] Mark E. Everett,[20] Eric B. Ford,[21] Jonathan Fortney,[8] Thomas N. Gautier III,[22] John C. Geary,[3] Alan Gould,[23] Michael Haas,[1] Christopher Henze,[1] Andrew W. Howard,[24] Steve B. Howell,[1] Daniel Huber,[1] Jon M. Jenkins,[5] Hans Kjeldsen,[18] Rea Kolbl,[6] Jeffery Kolodziejczak,[25] David W. Latham,[3] Brian L. Lee,[2] Eric Lopez,[8] Fergal Mullally,[5] Jerome A. Orosz,[26] Andrej Prsa,[27] Elisa V. Quintana,[5] Roberto Sanchis-Ojeda,[30] Dimitar Sasselov,[3] Shawn Seader,[5] Avi Shporer,[8,28] Jason H. Steffen,[29] Martin Still,[10] Peter Tenenbaum,[5] Susan E. Thompson,[5] Guillermo Torres,[3] Joseph D. Twicken,[5] William F. Welsh,[26] Joshua N. Winn[30]

We present the detection of five planets—Kepler-62b, c, d, e, and f—of size 1.31, 0.54, 1.95, 1.61 and 1.41 Earth radii ($R_\oplus$), orbiting a K2V star at periods of 5.7, 12.4, 18.2, 122.4, and 267.3 days, respectively. The outermost planets, Kepler-62e and -62f, are super–Earth-size (1.25 $R_\oplus$ < planet radius ≤ 2.0 $R_\oplus$) planets in the habitable zone of their host star, respectively receiving 1.2 ± 0.2 times and 0.41 ± 0.05 times the solar flux at Earth's orbit. Theoretical models of Kepler-62e and -62f for a stellar age of ~7 billion years suggest that both planets could be solid, either with a rocky composition or composed of mostly solid water in their bulk.

Kepler is a NASA Discovery-class mission designed to determine the frequency of Earth-radius planets in and near the habitable zone (HZ) of solar-like stars (*1–6*). Planets are detected as "transits" that cause the host star to appear periodically fainter when the planets pass in front of it along the observer's line of sight. Kepler-62 [Kepler Input Catalog (KIC) 9002278, Kepler Object of Interest (KOI) 701] is one of about 170,000 stars observed by the Kepler spacecraft. On the basis of an analysis of long-cadence photometric observations from Kepler taken in quarters 1 through 12 (13 May 2009 through 28 March 2012), we report the detection of five planets orbiting Kepler-62, including two super–Earth-size planets in the HZ as well as a hot Mars-size planet (Fig. 1 and Table 1). Before validation, three of these objects were designated as planetary candidates KOI-701.01, 701.02, and 701.03 in the Kepler 2011 catalog (*7*) and the Kepler 2012 catalog (*8*). KOI-701.04 and 701.05 were subsequently identified using a larger data sample (*9*).

Analysis of high-resolution spectra indicates that Kepler-62 is a K2V spectral type with an estimated mass and radius (in solar units) of 0.69 ± 0.02 $M_\odot$ and 0.63 ± 0.02 $R_\odot$ (*9*). Examination of the sky close to Kepler-62 showed the presence of only one additional star that contributed as much as 1% to the total flux (figs. S3 and S4) (*9*). Warm-Spitzer observations (fig. S9) and the analysis of centroid motion (table S1) were consistent with the target star as the source of the transit signals (Fig. 1 and fig. S1). We computed the radius, semimajor axis, and radiative equilibrium temperature of each planet (Table 1) on the basis of light curve modeling given the derived stellar parameters (table S3).

The masses of the planets could not be directly determined using radial velocity (RV) measurements of the host star because of the planets' low masses, the faintness and variability of the star, and the level of instrument noise. In the absence of a detected signal in the RV measurements (*9*), we statistically validated the planetary nature of Kepler-62b through -62f with the BLENDER procedure (*10–13*) by comparing the probability of eclipsing binaries and other false-positive scenarios to bona fide transiting planet signals (*14–18*).

To systematically explore the different types of false positives that can mimic the signals, we generated large numbers of synthetic light curves that blend together light from multiple stars and planets over a wide range of parameters, and then compared each blend with the Kepler photometry (Fig. 2). We rejected blends that resulted in light curves inconsistent with the observations.

We then estimated the frequency of the allowed blends by taking into account all available observational constraints from the follow-up observations discussed in (*9*). Finally, we compared this frequency with the expected frequency of true planets (planet "prior") to derive the odds ratio (*9*). By incorporating these constraints into a Monte Carlo model that considered a wide range of stellar and planetary characteristics, we determined estimates of the probability of each false positive that could explain the observations (*9*).

Our simulations of each of the candidates indicate that the likelihood of a false-positive explanation is much smaller than the likelihood that the candidates constitute a planetary system. The calculated odds ratios that Kepler-62b through -62f represent planets rather than false positives are 5400, >5000, 15,000, 14,700, and >5000, respectively (*9*). There is also a 0.2% chance that the planets orbit a widely spaced binary composed of two K2V stars; if so, the planets are larger in radius than the values shown in Table 1 by a factor of $\sqrt{2}$ (*9*).

To determine whether a planet is in the HZ, we calculated the flux of stellar radiation that it intercepts. It is convenient to express intercepted flux in units of the average solar flux intercepted by Earth, denoted by $S_\odot$. The values of the stellar flux intercepted by Kepler-62b to -62f are $70 \pm 9\ S_\odot$, $25 \pm 3\ S_\odot$, $15 \pm 2\ S_\odot$, $1.2 \pm 0.2\ S_\odot$, and $0.41 \pm 0.05\ S_\odot$, respectively. Eccentric planetary orbits increase the annually averaged irradiation from the primary star by a factor of $1/(1 - e^2)^{1/2}$, where e is the orbital eccentricity (*19*). Because the model results for the orbital eccentricities of Kepler-62b through -62f are small and consistent with zero, no corrections were made.

The HZ is defined here as the annulus around a star where a rocky planet with a CO_2-H_2O-N_2 atmosphere and sufficiently large water content (such as on Earth) can host liquid water on its solid surface (*20*). In this model, the locations of the two edges of the HZ are determined on the basis of the stellar flux intercepted by the planet and the assumed composition of the atmosphere. A conservative estimate of the range of the HZ (labeled "narrow HZ" in Fig. 3) is derived from atmospheric models by assuming that the planets have H_2O- and CO_2-dominated atmospheres with no cloud feedback (*21*). The flux range is defined at the inner edge by thermal runaway due to saturation of the atmosphere by water vapor and at the outer edge by the freeze-out of CO_2. In this model, the planets are assumed to be geologically active and climatic stability is provided by a mechanism in which atmospheric CO_2 concentration varies inversely with planetary surface temperature.

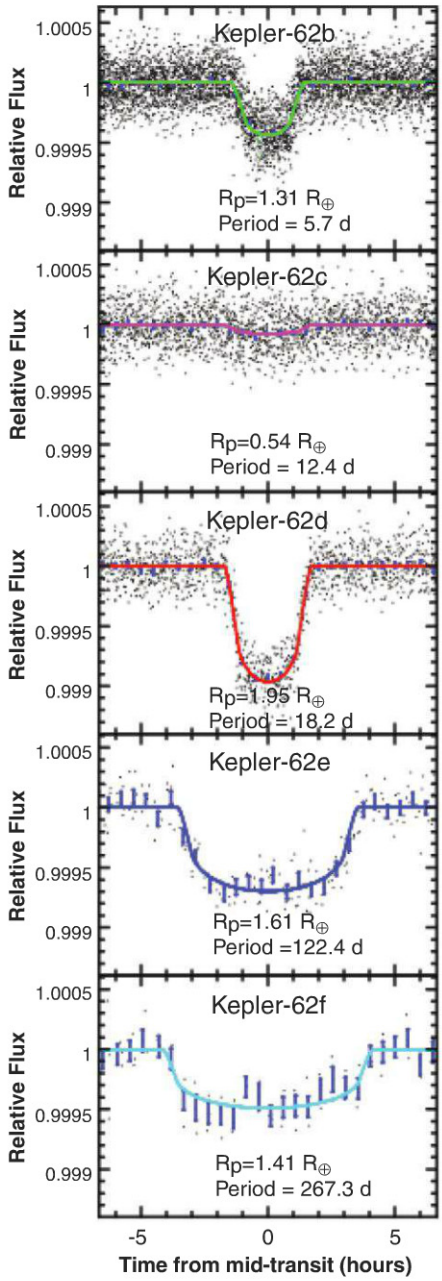


Fig. 1. Kepler-62 light curves after the data were detrended to remove the stellar variability. Composite of phase-folded transit light curves (dots), data binned in ½ hour intervals (blue error bars), and model fits (colored curves) for Kepler-62b through -62f. Model parameters are provided in Table 1. The error bars get larger as the period becomes larger because there are fewer points to bin together. For the shortest periods, the bars are too small to see.

The "empirical" HZ boundaries are defined by the solar flux received at the orbits of Venus and Mars at the epochs when they potentially had liquid water on their surfaces. Venus and Mars are believed to have lost their water at least 1 billion years and 3.8 billion years ago, respectively, when the Sun was less luminous. At these epochs, Venus received a flux of $1.78\ S_\odot$ and Mars a flux of $0.32\ S_\odot$ (*20*). The stellar spectral energy distributions of stars cooler than the Sun are expected to slightly increase the absorbed flux (*20*). Including this factor changes the HZ flux limits to 1.66 and $0.27\ S_\odot$ for the empirical HZ, and 0.95 and $0.29\ S_\odot$ for the narrow HZ (*21*). Figure 3 shows that Earth and Kepler-62f are within the flux boundaries of the narrow HZ, whereas Kepler-22b and Kepler-62e are within the empirical flux boundaries.

Although RV observations were not precise enough to measure masses for Kepler-62e and -62f, other exoplanets with a measured radius below $1.6\ R_\oplus$ have been found to have densities indicative of a rocky composition. In particular, Kepler-10b (*22*), Kepler-36b (*23*), and CoRoT-7b (*24*) have radii of $1.42\ R_\oplus$, $1.49\ R_\oplus$, and $1.58\ R_\oplus$ and densities of 8.8, 7.5, and 10.4 g/cm^3, respectively. Thus, it is possible that both Kepler-62e and -62f (with radii of $1.61\ R_\oplus$ and $1.41\ R_\oplus$) are also rocky planets.

The albedo and the atmospheric characteristics of these planets are unknown, and therefore the range of equilibrium temperatures (T_{eq}) at which the thermal radiation from each planet balances the insolation is large and depends strongly on the composition and circulation of the planets' atmospheres, their cloud characteristics and coverage, and the planets' rotation rates (*25*, *26*). However, for completeness, values of T_{eq} were computed from $T_{eq} = T_{eff}\,[\beta(1 - A_B)(R_*/2a)^2]^{1/4}$, where T_{eff} is the effective temperature of the star (4925 K), R_* is the radius of the star relative to the Sun (0.64), A_B is the planet Bond albedo, a is the planet semimajor axis, and β is a proxy for day-night redistribution (1 = full redistribution, 2 = no redistribution). For the Markov chain Monte Carlo calculations, it was assumed that $\beta = 1$ and that A_B is a random number from 0 to 0.5 (Table 1).

[1]NASA Ames Research Center, Moffett Field, CA 94035, USA. [2]Department of Astronomy, Box 351580, University of Washington, Seattle, WA 98195, USA. [3]Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA. [4]Max Planck Institute of Astronomy, Koenigstuhl 17, 69115 Heidelberg, Germany. [5]SETI Institute, Mountain View, CA 94043, USA. [6]University of California, Berkeley, CA 94720, USA. [7]Yale University, New Haven, CT 06520, USA. [8]Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064, USA. [9]Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA. [10]Bay Area Environmental Research Institute, Moffett Field, CA 94035, USA. [11]Vanderbilt University, Nashville, TN 37235, USA. [12]Carnegie Institution of Washington, Washington, DC 20015, USA. [13]McDonald Observatory, University of Texas, Austin, TX 78712, USA. [14]Niels Bohr Institute, University of Copenhagen, DK-2100 Copenhagen, Denmark. [15]Centre for Star and Planet Formation, Natural History Museum of Denmark, University of Copenhagen, DK-1350 Copenhagen, Denmark. [16]Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA. [17]Department of Physics, University of Notre Dame, Notre Dame, IN 46556, USA. [18]Department of Physics and Astronomy, Aarhus University, Aarhus, Denmark. [19]Exoplanet Science Institute/Caltech, Pasadena, CA 91125, USA. [20]National Optical Astronomy Observatory, Tucson, AZ 85719, USA. [21]University of Florida, Gainesville, FL 32611, USA. [22]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA. [23]Lawrence Hall of Science, Berkeley, CA 94720, USA. [24]Institute for Astronomy, University of Hawaii, Honolulu, HI 96822, USA. [25]Marshall Space Flight Center, Huntsville, AL 35805, USA. [26]San Diego State University, San Diego, CA 92182, USA. [27]Villanova University, Villanova, PA 19085, USA. [28]Las Cumbres Observatory Global Telescope, Goleta, CA 93117, USA. [29]Northwestern University, Evanston, IL 60208, USA. [30]Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

*Corresponding author. E-mail: william.j.borucki@nasa.gov

Gravitational interactions between Kepler-62e and -62f are too weak (*9*) to cause nonlinear variations in the times of transits (*27*, *28*) and thereby provide estimates of their masses. Nevertheless, upper limits (95th percentile) for Kepler-62e and -62f were derived (table S4): 150 $M_{\oplus}$ and 35 $M_{\oplus}$, respectively. The smallest upper limit to the mass of Kepler-62e based on RV observations (table S4) gives 36 $M_{\oplus}$. These values confirm their planetary nature without constraining their composition. Despite the lack of a measured mass for Kepler-62e and -62f, the precise knowledge of their radii, combined with estimates of their T_{eq} and the stellar age (~7 billion years), imply that Kepler-62e and -62f have lost their primordial or outgassed hydrogen envelope (*29*, *30*). Therefore, Kepler-62e and -62f are Kepler's first HZ planets that could plausibly be composed of

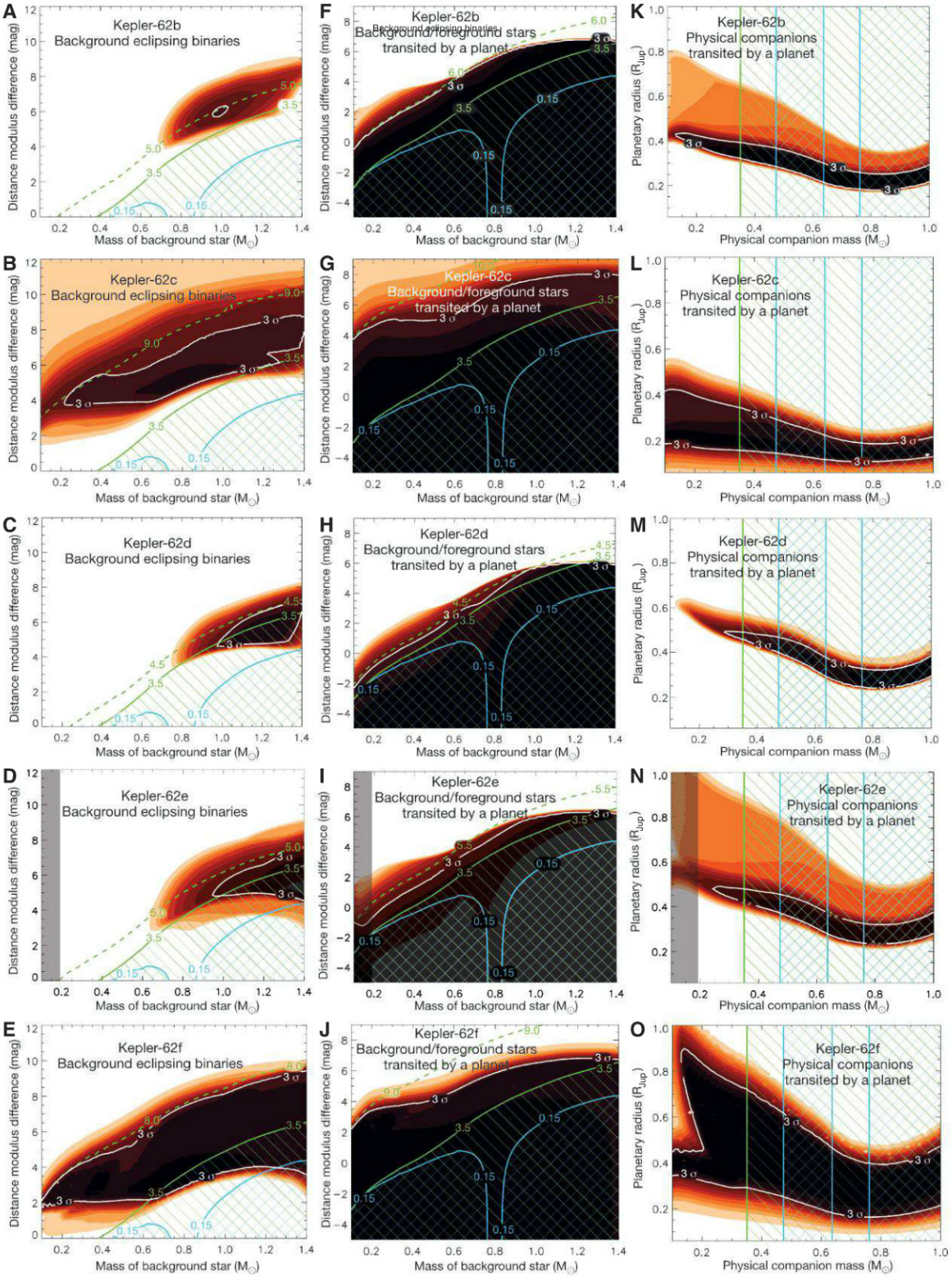


Fig. 2. BLENDER goodness-of-fit contours for Kepler-62b, c, d, e, and f corresponding to the three different scenarios that contribute to the overall blend frequency. (**A** to **E**) Background eclipsing binaries. (**F** to **J**) Background or foreground stars transited by a planet. (**K** to **O**) Physical companions transited by a planet. Viable blends must be less than ~5.0 magnitudes [(A) to (E)] or ~5.5 magnitudes [(F) to (J)] fainter than Kepler-62 (dashed line). Only blends inside the solid white contour match the Kepler light curve within acceptable limits (3σ, where σ is the significance level of the χ^2 difference compared to a transit model fit). Lighter-colored areas (red, orange, yellow) mark regions of parameter space giving increasingly worse fits to the data (4σ, 5σ, etc.) and correspond to blends that we consider to be ruled out. The cyan cross-hatched areas indicate regions of parameter space that we have ruled out because the resulting r-Ks color of the blend is either too red (left) or too blue (right), relative to the measured color, by more than 3σ (0.15 mag). The green hatched regions indicate blends that are ruled out because the intruding stars are less than 3.5 magnitudes fainter than the target and would be so bright that they would have been detected spectroscopically. Finally, the thin gray area at the left of (D), (I), and (N) rules out stars on the basis of our Spitzer observations (fig. S8) (*9*). The likelihood of a false positive for each planetary candidate is derived from the integration of the area that remains within the 3σ boundary that is not eliminated by the hatched areas.

Table 1. Characteristics of the Kepler-62 planetary system. T_0 is the epoch in mid-transit in barycentric Julian days, P is the period, "depth" is the percent reduction of the flux during the transits determined from the model fit to the data, R_p/R_* is the ratio of the radius of the planet to the radius of the star, a/R_* is the ratio of the planet's semimajor axis to the stellar radius, b is the impact parameter in units of stellar radius, i is the orbital inclination, $e \cos \omega$ is the product of the orbital eccentricity e and the cosine of the periapse angle ω, a is the planet semimajor axis, R_p is the planet radius, maximum mass is the upper limit to the mass based on transiting timing and RV observations, $M_\oplus$ is the mass of Earth, and T_{eq} is the radiative equilibrium temperature. The values of the uncertainties are ±1 standard deviation unless otherwise noted. Values for the maximum mass are for the 95th percentile (*9*). A second set of values for the planetary parameters was computed by an independent model and found to be in good agreement with the listed values.

Parameter	Kepler-62b	Kepler-62c	Kepler-62d	Kepler-62e	Kepler-62f
T_0 (BJD-2454900)	103.9189 ± 0.0009	67.651 ± 0.008	113.8117 ± 0.0008	83.404 ± 0.003	522.710 ± 0.006
P (days)	5.714932 ± 0.000009	12.4417 ± 0.0001	18.16406 ± 0.00002	122.3874 ± 0.0008	267.291 ± 0.005
Transit duration (hours)	2.31 ± 0.09	3.02 ± 0.09	2.97 ± 0.09	6.92 ± 0.16	7.46 ± 0.20
Depth (%)	0.043 ± 0.001	0.007 ± 0.001	0.092 ± 0.002	0.070 ± 0.003	0.042 ± 0.004
R_p/R_*	0.0188 ± 0.0003	0.0077 ± 0.0004	0.0278 ± 0.0006	0.0232 ± 0.0003	0.0203 ± 0.0008
a/R_*	18.7 ± 0.5	31.4 ± 0.8	40.4 ± 1.0	144 ± 4	243 ± 6
b	0.25 ± 0.13	0.16 ± 0.09	0.22 ± 0.13	0.06 ± 0.05	0.41 ± 0.14
i	89.2 ± 0.4	89.7 ± 0.2	89.7 ± 0.3	89.98 ± 0.02	89.90 ± 0.03
$e \cos \omega$	0.01 ± 0.17	−0.05 ± 0.14	−0.03 ± 0.24	0.05 ± 0.17	−0.05 ± 0.14
$e \sin \omega$	−0.07 ± 0.06	−0.18 ± 0.11	0.09 ± 0.09	−0.12 ± 0.02	−0.08 ± 0.10
a (AU)	0.0553 ± 0.0005	0.0929 ± 0.0009	0.120 ± 0.001	0.427 ± 0.004	0.718 ± 0.007
R_p ($R_\oplus$)	1.31 ± 0.04	0.54 ± 0.03	1.95 ± 0.07	1.61 ± 0.05	1.41 ± 0.07
Maximum mass ($M_\oplus$) (*9*)	9	4	14	36	35
Number of observed transits	171	76	52	8	3
Total SNR	54	8.5	68	31	12
T_{eq} (K)	750 ± 41	578 ± 31	510 ± 28	270 ± 15	208 ± 11

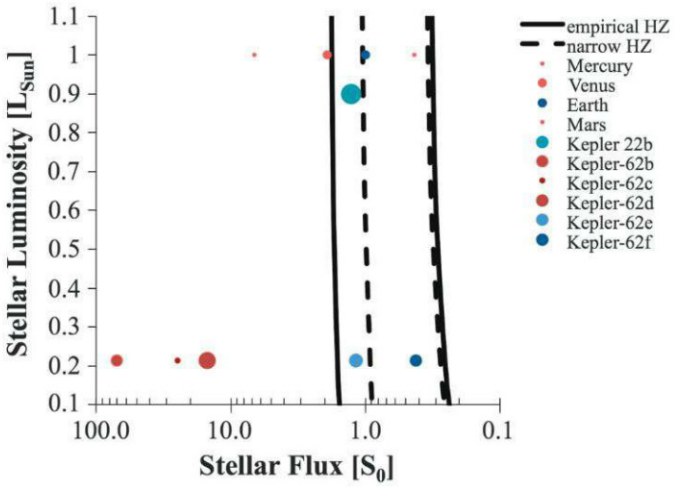


Fig. 3. Comparison of known HZ exoplanets with measured radii less than 2.5 $R_\oplus$ to the solar system planets. The sizes of the circles indicate the relative sizes of the planets to each other. The dashed and solid lines indicate the edges of the narrow and empirical HZ, respectively.

condensable compounds and be solid, either as a dry, rocky super-Earth or composed of a substantial amount of water (most of which would be in a solid phase because of the high internal pressure) surrounding a silicate-iron core.

We do not know whether Kepler-62e and -62f have a rocky composition, an atmosphere, or water. Until we get suitable spectra of their atmospheres, we cannot determine whether they are in fact habitable. With radii of 1.61 and 1.41 $R_\oplus$, respectively, Kepler-62e and -62f are the smallest transiting planets detected by the Kepler mission that orbit within the HZ of any star other than the Sun.

References and Notes

1. W. Borucki *et al.*, *Proc. IAU* **4**, S253 (2009).
2. N. M. Batalha *et al.*, *Astrophys. J.* **713**, L109 (2010).
3. D. G. Koch *et al.*, *Astrophys. J.* **713**, L79 (2010).
4. V. S. Argabright *et al.*, *Proc. SPIE* **7010**, 70102L (2008).
5. W. J. Borucki *et al.*, *Science* **327**, 977 (2010).
6. J. M. Jenkins *et al.*, *Astrophys. J.* **713**, L87 (2010).
7. W. J. Borucki *et al.*, *Astrophys. J.* **736**, 19 (2011).
8. N. M. Batalha *et al.*, http://arxiv.org/abs/1202.5852 (2012).
9. See supplementary materials on *Science* Online.
10. G. Torres, M. Konacki, D. D. Sasselov, S. Jha, *Astrophys. J.* **614**, 979 (2004).
11. G. Torres *et al.*, *Astrophys. J.* **700**, 1589 (2011).
12. F. Fressin *et al.*, *Astrophys. J. Suppl. Ser.* **197**, 5 (2011).
13. F. Fressin *et al.*, *Nature* **482**, 195 (2012).
14. J. M. Désert *et al.*, *Astron. Astrophys.* **526**, A12 (2011).
15. S. Ballard *et al.*, *Astrophys. J.* **743**, 200 (2011).
16. W. J. Borucki *et al.*, *Astrophys. J.* **745**, 120 (2012).
17. S. Howell *et al.*, *Astrophys. J.* **746**, 123 (2012).
18. T. N. Gautier III *et al.*, *Astrophys. J.* **749**, 15 (2012).
19. D. M. Williams, D. Pollard, *Int. J. Astrobiol.* **1**, 61 (2002).
20. J. F. Kasting, D. P. Whitmire, R. T. Reynolds, *Icarus* **101**, 108 (1993).
21. R. K. Kopparapu *et al.*, http://arxiv.org/abs/1301.6674 (2013).
22. N. M. Batalha *et al.*, *Astrophys. J.* **729**, 27 (2011).
23. J. A. Carter *et al.*, *Science* **337**, 556 (2012).
24. A. P. Hatzes *et al.*, *Astrophys. J.* **743**, 75 (2011).
25. F. Selsis *et al.*, *Astron. Astrophys.* **476**, 1373 (2007).
26. L. Kaltenegger, D. Sasselov, *Astrophys. J.* **736**, L25 (2011).
27. M. J. Holman, N. W. Murray, *Science* **307**, 1288 (2005).
28. E. Agol, J. Steffen, R. Sari, W. Clarkson, *Mon. Not. R. Astron. Soc.* **359**, 567 (2005).
29. L. Rogers, S. Seager, *Astrophys. J.* **716**, 1208 (2010).
30. E. D. Lopez, J. J. Fortney, N. K. Miller, http://arxiv.org/abs/1205.0010 (2012).

Acknowledgments: Kepler was competitively selected as the 10th Discovery mission. Funding for this mission is provided by NASA's Science Mission Directorate. Some of the data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership of the California Institute of Technology, the University of California, and NASA. The Keck Observatory was made possible by the generous financial support of the W. M. Keck Foundation. L.K. acknowledges support from Deutsche Forschungsgemeinschaft grants ENP Ka 3142/1-1 and NASA Astrobiology Institute. Funding for the Stellar Astrophysics Centre is provided by the Danish National Research Foundation. The research is supported by the ASTERISK project (Asteroseismic Investigations with SONG and Kepler) funded by European Research Council grant 267864. W.F.W. and J.A.O. acknowledge support from NASA through the Kepler Participating Scientist Program (PSP) and from NSF grant AST-1109928. D.F. acknowledges support from NASA ADAP12-0172. O.R.S.-O. and J.N.W. are supported by the Kepler PSP through grant NNX12AC76G. E.F. is partially supported by NASA PSP grants NNX08AR04G and NNX12AF73G. E.A. acknowledges NSF career grant AST-0645416. We also thank the Spitzer staff at IPAC and in particular N. Silbermann for checking and scheduling the Spitzer observations. The Spitzer Space Telescope is operated by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA. We thank the many people who gave so generously of their time to make this mission a success. All data products are available to the public at the Mikulski Archive for Space Telescopes, http://stdatu.stsci.edu/kepler.

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1234702/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S14
Tables S1 to S4
References (*31–57*)

2 January 2013; accepted 5 April 2013
Published online 18 April 2013;
10.1126/science.1234702

Complex N-Heterocycle Synthesis via Iron-Catalyzed, Direct C–H Bond Amination

Elisabeth T. Hennessy and Theodore A. Betley*

The manipulation of traditionally unreactive functional groups is of paramount importance in modern chemical synthesis. We have developed an iron-dipyrrinato catalyst that leverages the reactivity of iron-borne metal-ligand multiple bonds to promote the direct amination of aliphatic C–H bonds. Exposure of organic azides to the iron dipyrrinato catalyst furnishes saturated, cyclic amine products (N-heterocycles) bearing complex core-substitution patterns. This study highlights the development of C–H bond functionalization chemistry for the formation of saturated, cyclic amine products and should find broad application in the context of both pharmaceuticals and natural product synthesis.

Saturated, cyclic amines (N-heterocycles) are important building blocks for the synthesis of biologically active natural products, pharmaceutical agents, and materials. Current strategies for constructing saturated N-heterocycles are heavily dependent on functional group exchange, leading to inefficient synthetic protocols with poor atom economy and waste generation. A streamlined synthetic approach to this product class would rely on a catalyst capable of the direct amination of aliphatic C–H bonds. An advantage of this method is its potential to harness saturated hydrocarbon feedstocks. Unfortunately, current C–H bond functionalization protocols often require substrate preoxidation, directing groups, or strong chemical oxidants, which contribute to a lack of generality for this bond construction (*1*–*4*). Herein, we report an iron catalyst capable of functionalizing a broad range of aliphatic C–H bonds to form saturated, cyclic amine products.

A challenge to the development of a general and mild aliphatic C–H bond functionalization strategy is the unreactive nature of the substrates themselves. Saturated hydrocarbons are chemically inert due to the large C–H bond dissociation energy (93 to 105 kcal/mol) coupled with the energetic and spatial inaccessibility of the C–H bonding and antibonding orbitals. Nature provides a blueprint to overcome these obstacles. The reaction of dioxygen with heme iron in cytochrome P450 produces a strong oxidant consisting of an iron-oxygen multiple bond (iron-oxo) (*5*). The iron-oxo bond contains two electrons residing in Fe–O π* orbitals [$Fe(d_{xz},d_{yz}) - O(p_x,p_y)$], which result in a weakened Fe–O bond vector possessing radical character, thus rendering the entire unit a reactive functionality. As a consequence of this electronic configuration, the iron-oxo bond can activate substrate aliphatic C–H bonds via an H–atom abstraction mechanism and thereby circumvent the orbital spatial restrictions that hinder oxidative addition pathways. Subsequent substrate functionalization results from recombination of the organic radical generated in the activation step with the open-shell iron-hydroxyl to produce an alcohol product with concomitant reduction of the iron. Despite this 30-year-old mechanistic precedent (*6*), viable catalysts fashioned with these design principles are only now being discovered.

The direct functionalization of C–H bonds based on a strategy exemplified by cytochrome P450 would be transformative in converting ubiquitous C–H bonds into functional group handles and would circumvent the traditional synthetic requirement for functional group exchange (*7*). The electronic structure of the cytochrome P450 reactive iron-oxo intermediate can, in principle, be replicated with any metal-ligand multiple bond (*8*) and would constitute a general strategy for the conversion of unactivated C–H bonds into a variety of C–heteroatom bond products. Indeed, metal stabilized carbene and nitrene transfer has garnered considerable interest through the use of noble metal catalysts (*1*, *9*–*14*). Specifically, Fiori *et al.* (*12*) and Liang *et al.* (*13*) have developed a class of C–H amination Rh_2-dicarboxylate catalysts capable of generating cyclic carbamate, guanidine, and sulfamide products. Recently, this methodology has been extended to include aryl azides to produce indolines via an intramolecular sp^3 C–H amination, as reported by Nguyen *et al.* (*14*). In contrast, late, first-row transition metal complexes are potentially ideal catalyst candidates but have been less explored. Their high d-electron count and compressed ligand fields (compared with their second- and third-row analogs) favor population of metal-ligand antibonding orbitals leading to destabilization and reactivity akin to the cytochrome P450 iron-oxo intermediate (*15*–*24*). With these design principles in mind, we describe an iron-dipyrrinato catalyst that can selectively aminate sp^3 C–H bonds. Herein, we present the application of this catalyst toward the

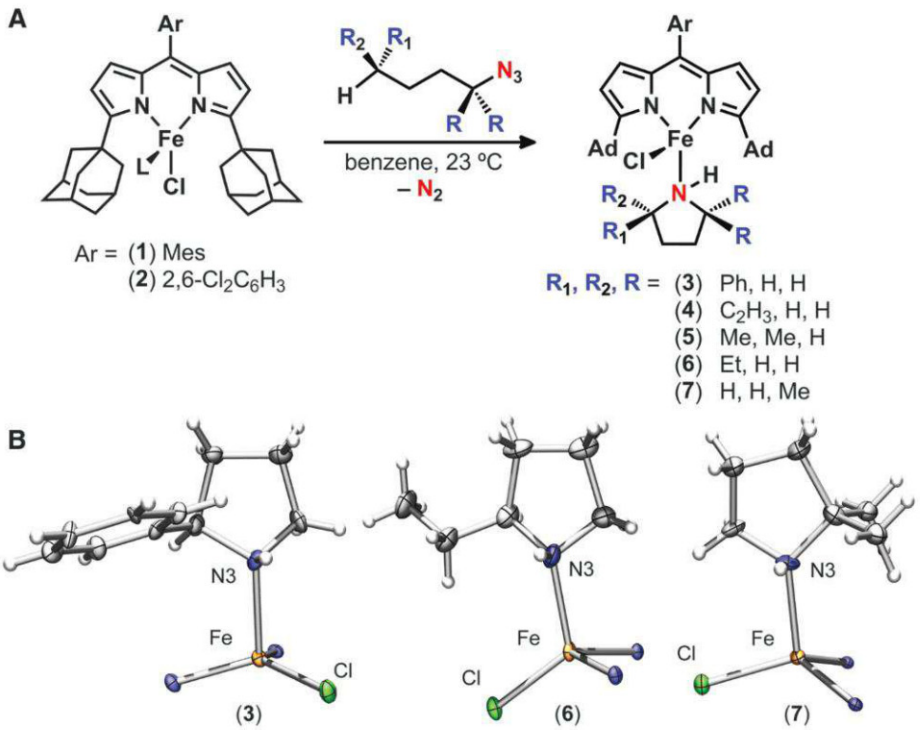


Fig. 1. Azide cyclization and iron-bound pyrrolidine products. (**A**) Reaction of complexes **1** and **2** with linear azides to generate Fe-bound pyrrolidine products **3** to **7**. (**B**) Solid-state core structures of (^{Ad}L)FeCl(2-Ph-NHC_4H_7) (**3**) from reaction with **1** and 1-azido-4-phenylbutane, (^{Ad}L)FeCl(2-Et-NHC_4H_7) (**6**) from reaction with **1** and 1-azidohexane, and (^{Ad}L)FeCl(2,2-Me_2-NHC_4H_6) (**7**) from reaction with **2** and 2-azido-2-methylpentane, with the thermal ellipsoids set at the 50% probability level (Fe, orange; C, gray; H, white; N, blue; Cl, green).

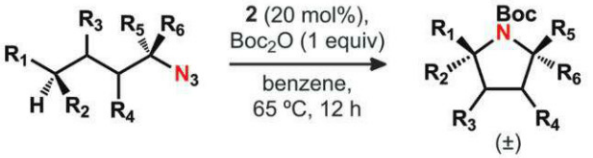
Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA.

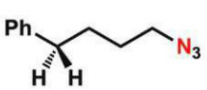
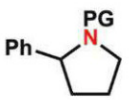
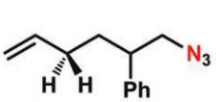
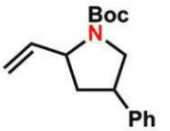
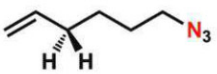
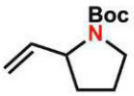
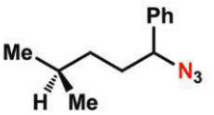
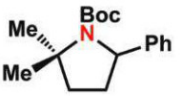
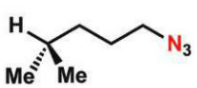
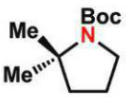
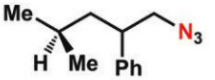
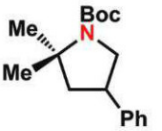
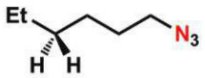
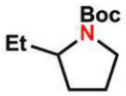
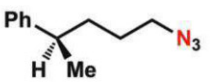
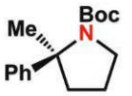
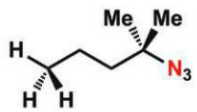
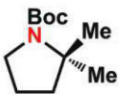
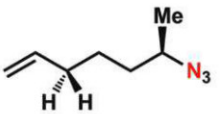
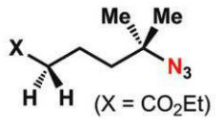
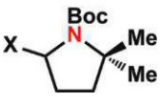
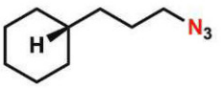
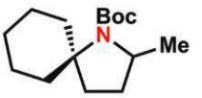
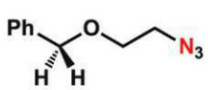
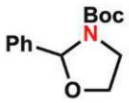
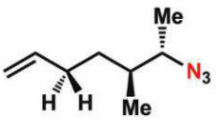
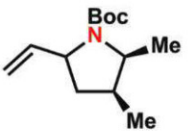
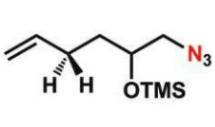
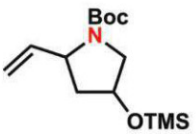
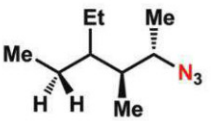
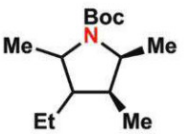
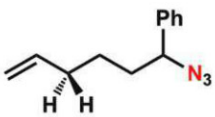
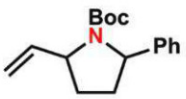
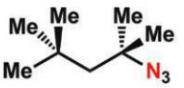
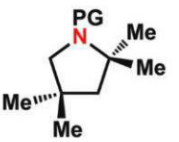
*Corresponding author. E-mail: betley@chemistry.harvard.edu

production of complex cyclic amine structures from simple linear aliphatic substrates requiring a single functionality, an azide (Fig. 1).

We previously reported intermolecular amination of benzylic C–H bonds with aryl and alkyl azides using the ferrous dipyrrinato complex (^{R}L)FeCl(solv) [R = 2,4,6-$Ph_3C_6H_2$, Ad with meso-Ar = mesityl (**1**), 2,6-$Cl_2C_6H_3$ (**2**); solv = Et_2O, tetrahydrofuran; L, ligand; Ph, phenyl; Ad, adamantyl; Ar, aryl; Et, ethyl] (*19*). Isolation and characterization of the reactive intermediate elucidated the electronic structure of the high-spin, iron-bound imido radical, wherein a high-spin Fe(III) ($S = {}^{5}/_{2}$) is antiferromagnetically coupled to the imido radical ($S = -{}^{1}/_{2}$) to give a high-spin ground state. This electronic structure places substantial radical character on both the Fe–N σ and π bond vectors, facilitating both radical

Table 1. Catalytic synthesis of pyrrolidine products. R_1: Ph, $CHCH_2$, Me, Et, H, CO_2Et, $(CH_2)_5$; R_2: H, Me, $(CH_2)_5$; R_3: H, Et, Me; R_4: H, OTMS (trimethylsiloxy), Ph, Me; R_5: H, Me, Ph; R_6: H, Me. dr, diastereomeric ratio.



Entry	Azide	Pyrrolidine	Yield (%)*	Entry	Azide	Pyrrolidine	Yield (%)*
1			98†§ (PG = Fmoc) 93†§ 57‡§ (PG = Boc)	10			66 1:5:1.0 dr
2			72‡§	11			70
3			60§ 49‡§	12			98
4			19‡§	13			75‡ 93% ee
5			17‡§	14			84 1.1:1.0 dr
6	 (X = CO_2Et)		11	15			67
7			47	16			73 2.1:1.0 dr
8			68‡	17			58 5.5:1.5: 1.0:0.08 dr
9			60 3.9:1.0 dr	18			14 (PG = Boc) 78† (PG = Fmoc)

* ^{1}H NMR yield using ferrocene or trimethoxybenzene as an internal standard unless otherwise noted. † Stoichiometric reaction with one equivalent of catalyst **2**. ‡ Isolated yield. § 10 mol% catalyst **2**.

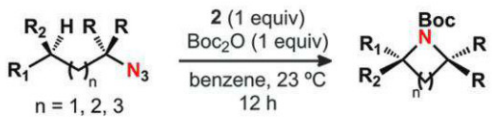
H-atom abstraction and radical recombination pathways to proceed. Furthermore, the amination catalytic cycle remains in the quintet spin state ($S = 2$), making each step of the catalytic cycle spin-allowed. We envisioned that the intramolecular extension of this reactivity profile to linear, aliphatic azide substrates would generate cyclic amine products in a single operation.

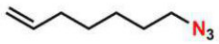
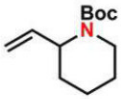
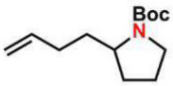
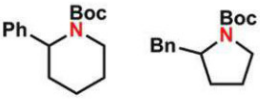
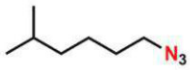
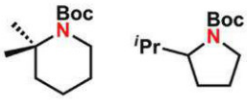
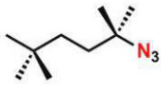
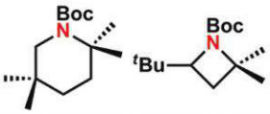
To test the viability of catalyst **1** for intramolecular C–H amination, we subjected a variety of substituted aliphatic azides to complex **1** (Fig. 1A). Exposure of 1-azido-4-phenylbutane to **1** at room temperature in benzene resulted in consumption of the azide, as ascertained by the disappearance of the azide stretch in the infrared spectrum, and afforded a new paramagnetically shifted ^{1}H nuclear magnetic resonance (NMR) spectrum. Crystallization occurred from a concentrated hexanes solution of the product at 23°C to yield crystals in which 2-phenylpyrrolidine was bound to the (^{Ad}L)FeCl complex (**3**, Fig. 1B) (*25*). Similarly, treatment of 1-azido-5-hexene with **1** afforded the cyclized product 2-vinylpyrrolidine as an iron-bound adduct (**4**, fig. S5). In addition to allylic and benzylic C–H bonds, less reactive tertiary C–H bonds could be similarly functionalized. The reaction of 1-azido-5-methylpentane and **1** under standard conditions gave the 2,2-dimethylpyrrolidine iron-bound product (**5**). Gratifyingly, even secondary aliphatic C–H bonds could be functionalized through this method. Addition of 1-azidohexane to **1** resulted in the rapid consumption of the azide to afford the 2-ethylpyrrolidine complex (**6**, Fig. 1B). In an attempt to activate the primary C–H bond of an aliphatic azide substrate, 1-azidobutane was exposed to **1**. However, the only products observed in this transformation were linear *n*-butylamine and *n*-butylimine. To eliminate the potential for imine formation, through a process involving intermolecular C–H bond activation or β-hydride elimination, the gem-dimethyl substrate 2-azido-2-methylpentane was prepared and subjected to **1** at room temperature for 6 hours to afford the cyclized 2,2-dimethylpyrrolidine complex (**7**, Fig. 1B) in quantitative yield. The presence of the two α-Me (Me, methyl) substituents may facilitate the C–H bond functionalization and cyclization process through the Thorpe-Ingold effect (*26*).

With a reliable protocol for the stoichiometric C–H functionalization of aliphatic azides in hand, we attempted to render the reaction catalytic [5 to 10 equivalents (equiv.) of azide per 1 equiv. of **1** or **2**]. Unfortunately, examination of the cyclization reaction under catalytic conditions did not markedly increase the yield of the resultant free heterocyclic product. We attribute the lack of catalyst turnover to product inhibition, in which a tight Lewis acid/base pair between the dipyrrinato iron and the heterocyclic nitrogen atom is formed (Fig. 1B). To overcome product inhibition, we performed the cyclization reaction in the presence of an in situ protection reagent, which would reduce the nucleophilicity of the product N-heterocycle while avoiding the generation of by-products that might retard or prevent catalysis. Accordingly, treatment of a solution of 1-azido-4-phenylbutane and 9-fluorenylmethyl *N*-succinimidyl carbonate (Fmoc-OSuc; a base-cleavable protecting group) in benzene at room temperature with a stoichiometric amount of **2** for 12 hours afforded the Fmoc-protected 2-phenylpyrrolidine in 98% yield (Table 1, entry 1). Similarly, addition of an equivalent of **2** to a solution of 1-azido-4-phenylbutane and di-*tert*-butyl dicarbonate (Boc_2O; an acid-cleavable protecting group) under similar reaction conditions afforded the *tert*-butyloxycarbonyl (Boc)–protected product 1-Boc-2-phenylpyrrolidine in 93% yield. As catalyst loading was decreased, the *N*-hydroxysuccinimide by-product of Fmoc-protection led to catalyst decomposition through ligand protonation and limited the reaction to a single turnover. Fortunately, the by-products of protection with Boc_2O (^{t}BuOH, CO_2; ^{t}Bu, *tert*-butyl) did not inhibit catalyst turnover, permitting the heterocycle to be synthesized with catalytic amounts of **2** (chosen to eliminate benzylic C–H bonds from catalyst *meso*-aryl substituent).

We investigated application of catalytic quantities of **2** to the established in situ protection protocol for C–H functionalization and cyclization (Table 1). Exposure of substrates containing allylic, benzylic, or tertiary C–H bonds to **2** (10 mol %) provided the corresponding Boc-protected pyrrolidine in good isolated yield (49 to 72%, entries 1 to 3). Catalytic functionalization of a secondary C–H bond was also possible, affording 1-Boc-2-ethylpyrrolidine in a lower isolated yield (19%, entry 4). The formation of undesired linear side products is competitive with cyclization at the strong secondary C–H bond and contributes to the low yield. Even a primary C–H bond in 2-azido-2-methylpentane could be functionalized to give 1-Boc-2,2-dimethylpyrrolidine in 17% isolated yield (entry 5), with a mass balance of unreacted azide. Unlike the previous example, we hypothesize that the diminished yield reflects the ability of the tertiary azide to access the iron catalyst. In situ ^{1}H NMR monitoring of the stoichiometric reaction between **1** and 2-azido-2-methylpentane to generate the 2,2-dimethylpyrrolidine iron adduct requires 55 min at room temperature (with no detectable buildup of an intermediate), whereas conversion of 1-azido-5-methylpentane to the same product is complete in 5 min (figs. S1 and S2). The long reaction time to aminate the primary C–H bond likely has an adverse effect on the overall catalysis, as competitive catalyst inactivation occurs on a similar time scale. We then expanded the substrate scope to include hetero-atom–containing functional groups. Exposure of ethyl-5-azidopentanoate to **2** under standard catalytic conditions resulted in only linear primary amine and imine products. Again, blocking the α position of the azide with gem-dimethyl substituents led to productive cyclization (entry 6), albeit in the low yield characteristic of tertiary azide substrates. Introduction of heteroatoms between the reactive functionalities allowed for the formation of

Table 2. Product distribution for azetidine, pyrrolidine, and piperidine formation. R: H, Me; R_1: $CHCH_2$, $(CH_2)_2CHCH_2$, Ph, Me, CMe_3; R_2: H, Me; GC/MS, gas chromatography–mass spectrometry.



Entry	Azide	Product(s)	Conv. (%)*†
1			45
2			82
3			52 (1.0:0.9)
4			47 (1.0:1.5)
5			47 (1.0:1.5)

* Yields determined by ^{1}H NMR using ferrocene or trimethoxybenzene as internal standards. † Ratios determined by integration of GC/MS peaks.

1-Boc-2-phenyloxazolidine in 47% yield (entry 7). The reaction is also tolerant of siloxy groups, as shown in the formation of 1-Boc-2-vinyl-4-trimethylsiloxypyrrolidine in 68% isolated yield (entry 8).

Next, we explored the synthesis of highly substituted pyrrolidine products using substrates accessible via cuprate-assisted epoxide opening (*27*) followed by azide formation (Table 1, entries 9 to 18). The use of commercially available epoxide and alkyl halide building blocks permits virtually any substitution pattern to be programmed into the ensuing heterocyclic product. After $Li_2[CuCl_4]$ promoted epoxide opening, the resultant primary or secondary alcohols were tosylated and displaced with sodium azide to provide the desired cyclization precursor. Alternatively, tertiary and benzylic alcohols were directly converted to the corresponding azide by exposure to trimethylsilylazide and boron trifluoride diethyl etherate (*28*). Allylic, tertiary, and secondary C–H bond substrates available through either reaction sequence underwent facile C–H functionalization and cyclization (entries 9 to 17, 58 to 98% yield). Notably, this method provides access to an all-carbon spiro-center (entry 15, 67%). Functionalization of primary C–H bonds required stoichiometric catalyst loading (entry 18, 78% yield), for reasons stated above. Use of (*R*)-2-phenyl-5-azidopentane [95% enantiomeric excess (ee)] in the catalytic transformation resulted in (*S*)-2-methyl-2-phenylpyrrolidine with retention of configuration (entry 13, 75%, 93% ee). Application of stoichiometric quantities of catalyst **1** to the reaction gave the corresponding (*S*)-2-methyl-2-phenylpyrrolidine iron-bound adduct whose absolute stereochemistry was verified by x-ray diffraction (fig. S8).

Last, we investigated the potential of this method to generate N-heterocycles of various ring sizes. We anticipated that a vinyl directing group could be employed to encourage the site-selective functionalization of the allylic C–H bond within the acyclic precursor. Gratifyingly, treatment of 1-azido-6-heptene and Boc_2O (1 equiv.) with **2** (1 equiv.) at room temperature generated the six-membered 1-Boc-2-vinylpiperidine (Table 2, entry 1) as the exclusive reaction product. Use of the vinyl activating group to target seven-membered azepane products, however, led to exclusive formation of the corresponding pyrrolidine (entry 2). In contrast, a phenyl activating group was not effective in favoring the formation of a six-membered-ring product. Addition of **2** to 1-azido-5-phenyl-pentane under standard conditions resulted in a 1:0.85 mixture of both 1-Boc-2-phenylpiperidine and 1-Boc-2-benzylpyrrolidine (entry 3). Similarly, the use of a tertiary C–H bond to favor six-membered-ring formation in the case of 1-azido-5-methylhexane resulted in a mixture of piperidine and pyrrolidine products (entry 4). An attempt was made to block the potential for pyrrolidine formation; exposure of 2-azido-2,5,5-trimethylhexane and Boc_2O to **2** resulted in both the anticipated 1-Boc-2,2,5,5-tetramethylpiperidine and the unexpected 1-Boc-2,2-dimethyl-4-*tert*-butylazetidine (entry 5). Alternatively, use of catalyst **1** and omission of Boc_2O allowed for the characterization of the corresponding iron-bound piperidine and azetidine adducts by x-ray analysis (figs. S10 and S11).

As illustrated in Fig. 2A, we hypothesize that the C–H bond functionalization reaction occurs via a three-step process involving: (i) oxidation of the Fe^{II} catalyst to an Fe^{III} imido radical by the alkyl azide substrate, (ii) intramolecular H-atom abstraction to generate an alkyl radical and an Fe^{III} amide (path Ia), and (iii) radical recombination to form the observed N-heterocyclic product (path Ib). Alternatively, a direct C–H bond insertion by the Fe^{III} imido radical intermediate cannot be excluded (path II). Both mechanisms require that the substrate C–H bond be brought into close proximity to the reactive Fe-imido radical. Based on our previous findings, the imido radical likely resides in the plane defined by the iron and the dipyrrin ligand, flanked by large pyrrolide adamantyl substituents. Such a conformation requires that the C–H bond substrate approach the imido radical opposite the chloride ligand. We expect that once this orientation is obtained, C–H bond functionalization is rapid. This hypothesis is supported by the retention of stereochemical information during the cyclization of (*R*)-2-phenyl-5-azidopentane (Fig. 2B). This stereoretention probably reflects the spatial constraints imposed by the ligand adamantyl units to inhibit racemization of the carboradical intermediate. Additionally, cyclization of 1-azido-4-deutero-4-phenylbutane provides an intramolecular kinetic isotope effect of 5.3 at 25°C and 5.1(2) at 65°C (Fig. 2C). This value is similar to the kinetic isotope effect (KIE) observed in the hydroxylation of 1,3-dideuteroadamantane catalyzed by tetramesitylporphyrin iron with oxone [KIE = 4.1(2)] (*29*). Finally, addition of the radical clock substrate (2-(4-azidobutyl)cyclopropyl) benzene to **2** exclusively furnishes the pyrrolidine product 1-Boc-2-(2-phenylcyclopropyl)pyrrolidine with the cyclopropyl unit intact (Fig. 2D). The nonunity intramolecular kinetic isotope effect suggests a stepwise mechanism for benzylic substrates (Fig. 2A, path I), which is consistent with our previously reported intermolecular amination reaction (*19*). The stereospecificity of the cyclization and the preservation of the cyclopropyl unit in the radical clock experiment suggest that if a stepwise mechanism is operative, the radical intermediate following H-atom abstraction is short-lived [recombination rate > 10^{11} s^{-1} (*30*)]. Alternatively, the reaction mechanism may change to a direct-insertion pathway (Fig. 2A, path II) when stronger substrate C–H bonds are functionalized.

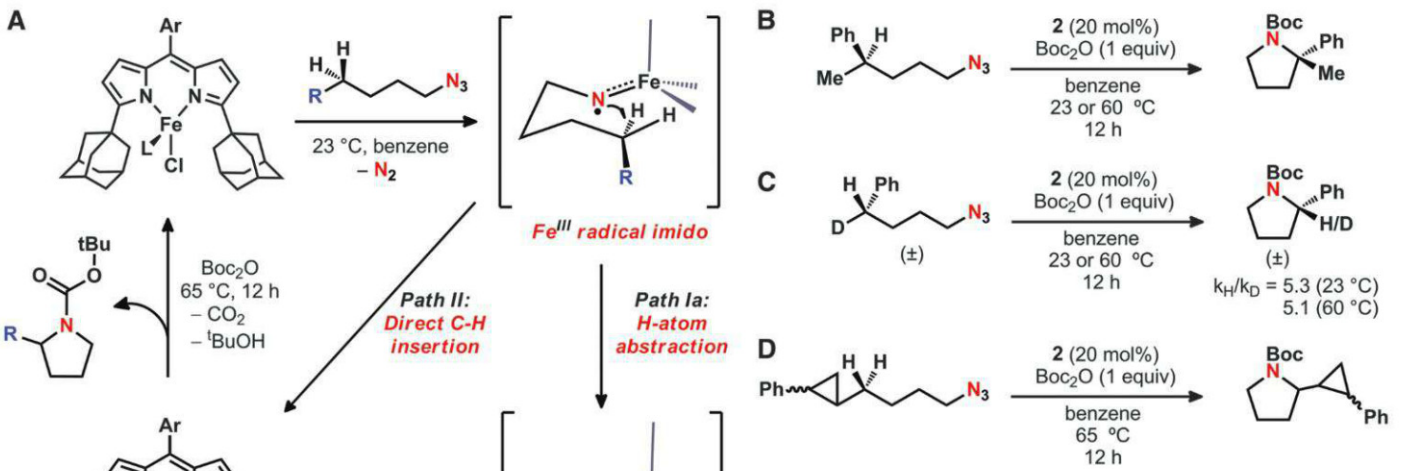


Fig. 2. Mechanistic studies. (**A**) Proposed mechanistic pathways for intramolecular C–H amination of linear alkyl azides with Fe catalyst **1** or **2** to form N-heterocycles. Pyrrolidine formation is depicted, although azetidine and piperidine products are also accessible. (**B** to **D**) Substrates designed to probe the mechanism of C–H functionalization and distinguish between paths I and II.

The foregoing results have demonstrated the oxidative potency of the transiently formed, high-spin iron imido radical for the functionalization of both activated and unactivated aliphatic C–H bond substrates. This iron-mediated cyclization of linear azides provides facile entry into complex N-heterocyclic products from readily available substrates that cannot be achieved by azide photolysis (*31*) or via classic Hoffmann-Löffler-Freytag methodologies (*32*). We anticipate the methodology described herein can be extended to produce a wide variety of saturated, cyclic structures.

References and Notes

1. D. N. Zalatan, J. Du Bois, *Top. Curr. Chem.* **292**, 347 (2010).
2. A. D. Ryabov, *Chem. Rev.* **90**, 403 (1990).
3. A. R. Dick, M. S. Sanford, *Tetrahedron* **62**, 2439 (2006).
4. J. A. Labinger, J. E. Bercaw, *Nature* **417**, 507 (2002).
5. P. R. Ortiz de Montellano, Ed., *Cytochrome P450: Structure, Mechanism, and Biochemistry* (Kluwer Academic/Plenum, New York, ed. 4, 2005).
6. J. T. Groves, G. A. McClusky, R. E. White, M. J. Coon, *Biochem. Biophys. Res. Commun.* **81**, 154 (1978).
7. A. O. King, N. Yasuda, *Top. Organometallic Chem.* **6**, 205 (2004).
8. W. A. Nugent, J. M. Mayer, *Metal-Ligand Multiple Bonds: The Chemistry of Transition Metal Complexes Containing Oxo, Nitrido, Imido, Alkylidene, or Alkylidyne Ligands* (Wiley, New York, 1988).
9. S. M. Au, J. S. Huang, W. Y. Yu, W. H. Fung, C. M. Che, *J. Am. Chem. Soc.* **121**, 9120 (1999).
10. H. M. L. Davies, J. R. Manning, *Nature* **451**, 417 (2008).
11. T. G. Driver, *Org. Biomol. Chem.* **8**, 3831 (2010).
12. K. W. Fiori, C. G. Espino, B. H. Brodsky, J. Du Bois, *Tetrahedron* **65**, 3042 (2009).
13. C. Liang *et al.*, *J. Am. Chem. Soc.* **130**, 343 (2008).
14. Q. Nguyen, K. Sun, T. G. Driver, *J. Am. Chem. Soc.* **134**, 7262 (2012).
15. L. Que Jr., *Acc. Chem. Res.* **40**, 493 (2007).
16. M. S. Chen, M. C. White, *Science* **318**, 783 (2007).
17. Y. M. Badiei *et al.*, *Angew. Chem. Int. Ed.* **47**, 9961 (2008).
18. C. A. Laskowski, A. J. M. Miller, G. L. Hillhouse, T. R. Cundari, *J. Am. Chem. Soc.* **133**, 771 (2011).
19. E. R. King, E. T. Hennessy, T. A. Betley, *J. Am. Chem. Soc.* **133**, 4917 (2011).
20. V. Lyaskovskyy *et al.*, *J. Am. Chem. Soc.* **133**, 12264 (2011).
21. S. Wiese *et al.*, *J. Am. Chem. Soc.* **134**, 10114 (2012).
22. S. M. Paradine, M. C. White, *J. Am. Chem. Soc.* **134**, 2036 (2012).
23. J. T. Groves, *J. Inorg. Biochem.* **100**, 434 (2006).
24. Q. Nguyen, T. Nguyen, T. G. Driver, *J. Am. Chem. Soc.* **135**, 620 (2013).
25. Materials, methods, and x-ray diffraction details are available as supplementary materials on *Science* Online.
26. R. M. Beesley, C. K. Ingold, J. F. Thorpe, *J. Chem. Soc. Trans.* **107**, 1080 (1915).
27. J. Drouin, F. Leyendecker, J. M. Conia, *Tetrahedron* **36**, 1195 (1980).
28. T. Mukaiyama, K. Kuroda, Y. Maruyama, *Heterocycles* **80**, 63 (2010).
29. A. Sorokin, A. Robert, B. Meunier, *J. Am. Chem. Soc.* **115**, 7293 (1993).
30. M. Newcomb, P. H. Toy, *Acc. Chem. Res.* **33**, 449 (2000).
31. D. H. R. Barton, A. N. Starratt, *J. Chem. Soc.* 2444 (1965).
32. A. W. Hoffman, *Berichte* **18**, 109 (1885).

Acknolwedgments: This work was supported by a grant from the NSF (CHE-0955885) and Harvard University. E.T.H. is grateful for a predoctoral fellowship from the U.S. Department of Energy Office of Science Graduate Fellowship Program (DE-AC05-060R23100) administered by Oak Ridge Institute for Science Education–Oak Ridge Associated Universities. T.A.B. is grateful for a George W. Merck Fellowship. We thank S.-L. Zheng for assistance with crystallography; J. Hong, A. Saghatelian, and B. Milgram for helpful discussions; and E. N. Jacobsen and T. Ritter for the generous use of their high-performance liquid chromatography and GC/MS instruments, respectively. The crystallographic data CCDC-905454 – 905461 can be obtained free of charge from the Cambridge Crystallographic Data Centre (www.ccdc.cam.ac.uk/data_request/cif).

Supplementary Materials

www.sciencemag.org/cgi/content/full/340/6132/591/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S11
Table S1
References (*33*–*52*)

5 December 2012; accepted 12 March 2013
10.1126/science.1233701

Self-Assembling Cages from Coiled-Coil Peptide Modules

Jordan M. Fletcher,[1] Robert L. Harniman,[1,2,3] Frederick R. H. Barnes,[1] Aimee L. Boyle,[1] Andrew Collins,[1,4] Judith Mantell,[5] Thomas H. Sharp,[1,5]* Massimo Antognozzi,[2,3] Paula J. Booth,[5] Noah Linden,[6] Mervyn J. Miles,[2,3] Richard B. Sessions,[5] Paul Verkade,[5,7] Derek N. Woolfson[1,5]†

An ability to mimic the boundaries of biological compartments would improve our understanding of self-assembly and provide routes to new materials for the delivery of drugs and biologicals and the development of protocells. We show that short designed peptides can be combined to form unilamellar spheres approximately 100 nanometers in diameter. The design comprises two, noncovalent, heterodimeric and homotrimeric coiled-coil bundles. These are joined back to back to render two complementary hubs, which when mixed form hexagonal networks that close to form cages. This design strategy offers control over chemistry, self-assembly, reversibility, and size of such particles.

In viruses (*1*) and certain bacterial microcompartments (*2*), capsids and suprastructures are produced via the self-assembly of large folded proteins, usually in highly symmetric manners and with exquisite positioning of noncovalent protein-protein interactions. Biomimetic assemblies have potential for creating simpler encapsulation systems and for applications in controlled delivery and release, sensing, and the preparation of protocells for various aspects of synthetic biology (*3*–*5*). To these ends, others have produced macroscopic "sacs" from peptide amphiphiles (*6*) and engineered micelle-like structures (*7*, *8*), small polyhedra (*9*–*11*), extended protein arrays (*12*), and metal-directed assemblies (*13*, *14*) using mainly natural peptides and proteins. Here, we show that self-assembled cage-like particles, SAGEs, can be made from a set of short, de novo, α-helical, coiled-coil peptides by using clear sequence-to-structure relationships and rational-design principles so as to direct stable and highly specific protein-protein interactions (*15*–*17*).

Previously, we developed a toolkit of coiled coils comprising homo-dimer, -trimer, and -tetramers (*18*) and a number of heterodimers (*19*). These synthetic peptides, of ≈30 residues in length, assemble reversibly and form stable structures at micromolar to nanomolar concentrations. To expand this toolkit and to ease the construction of the building blocks for the SAGE design, we engineered two new coiled-coil modules: (i) a shorter (~20 residues) homotrimer (CC-Tri3) and (ii) a similarly short obligate heterodimer (CC-Di-AB) comprising acidic (CC-Di-A) and basic (CC-Di-B) sequences (fig. S1 and table S1) (*20*). We chose a heterodimer for the second module in order to give control in the following self-assembly process. Our goal was to link copies of CC-Tri3 and CC-Di-A or CC-Di-B through their external surfaces via disulfide bonds (Fig. 1). These covalent constructs, dubbed CC-Tri3–CC-Di-A and CC-Tri3–CCDi-B, should assemble into complementary trimeric hubs, hub A and hub B, respectively. Alone, these should be water-soluble, discrete, partly folded helical structures; CC-Tri3 should spontaneously assemble, leaving CC-Di-A and CC-Di-B orphaned on the outside of the assemblies. Upon mixing, however, the two hubs should coassemble via association of the CC-Di-A and CC-Di-B modules to produce hexagonal networks with pores of ≈5 to 6 nm. Because the hubs are flexible and to maximize coiled-coil interactions, we argue that these networks should fold to form closed objects—SAGEs.

The two coiled coils were synthesized (fig. S2) and characterized in solution by using a combination of circular dichroism (CD) spectroscopy to measure secondary structure, stability, and dissociation constants (K_d values) (Fig. 2A, figs. S3 and S4, and tables S2 and S3); dynamic light scattering (DLS) (Fig. 2B and fig. S5); and analytical ultracentrifugation (AUC) (fig. S6) to probe peptide association. These methods confirmed

[1]School of Chemistry, Cantock's Close, University of Bristol, Bristol BS8 1TS, UK. [2]School of Physics, H. H. Wills Physics Laboratory, Tyndall Avenue, University of Bristol, Bristol BS8 1TL, UK. [3]Centre for Nanoscience and Quantum Information (NSQI), Tyndall Avenue, University of Bristol, Bristol BS8 1FD, UK. [4]Bristol Centre for Functional Nanomaterials, NSQI, Tyndall Avenue, University of Bristol, Bristol BS8 1FD, UK. [5]School of Biochemistry, Medical Sciences Building, University Walk, University of Bristol, Bristol BS8 1TD, UK. [6]School of Mathematics, University Walk, University of Bristol, Bristol BS8 1TW, UK. [7]School of Physiology and Pharmacology, Medical Sciences Building, University Walk, University of Bristol, Bristol BS8 1TD, UK.

*Present address: Department of Physics, Clarendon Laboratory, Parks Road, University of Oxford, Oxford OX1 3PU, UK.
†Corresponding author. E-mail: d.n.woolfson@bristol.ac.uk

CC-Tri3 as a highly helical trimeric assembly, with concentration-dependent folding (K_d, 20°C = 3.99×10^{-14} M^2), and a midpoint of thermal unfolding (T_M) of 65°C at 50 μM peptide. Similarly, CC-Di-A and CC-Di-B alone were unfolded in the micromolar range, but coassembled when mixed to form a helical heterodimer, CC-Di-AB, (K_d, 20°C = 5.83×10^{-8} M; T_M = 51°C) (Table 1). We verified that CC-Tri3 and CC-Di-AB did not form mixed species in the presence of each other by showing that the melting profile of the two coiled coils, when mixed, was the same as the average of the two independent profiles (fig. S7).

Building toward hubs A and B, the two-peptide constructs CC-Tri3–CC-Di-A and CC-Tri3–CC-Di-B had reduced mean residue ellipticities (MREs) compared with CC-Tri3 alone. Moreover, these values were close to averages of CC-Tri3 plus either CC-Di-A or CC-Di-B, respectively (Fig. 2A, Table 1, and fig. S3). In addition, the melting curves for the hubs were near simple averages of the component curves (figs. S8 and S9). Next, we mixed three equivalents of CC-Di-A with hub B, and of CC-Di-B with hub A—equimolar amounts of the underlying peptide components CC-Di-A and CC-Tri3–CC-Di-B, and of CC-Di-B and CC-Tri3–CC-Di-A. In both cases, this should produce "terminated," nine-helix assemblies (Fig. 1). Indeed, the increased MREs observed were indicative of near-complete folding of all of the modules (Fig. 2A, Table 1, and fig. S3). Moreover, the thermal denaturation curves for these assemblies were sigmoidal, and the apparent T_M values measured were near the theoretical value for fully decoupled folding of the CC-Tri3 and CC-Di-AB components of 55°C (Table 1, fig. S9, and table S2). In all of these cases, DLS showed that the particle sizes of the peptide modules, hubs, and terminated assemblies were ≈2 to 5 nm, which is consistent with discrete and appropriately sized objects (Fig. 2B). AUC (fig. S6) gave solution molecular weights consistent with the compositions of each of the assemblies (Table 1), except for the terminated hub B, which had a mass higher than expected but nonetheless was still a discrete assembly.

These findings all corroborate the modular design approach that underpins the SAGE concept.

When hub A and hub B were mixed in an equimolar ratio, a fine white precipitate formed within minutes, accounting for the >90% of peptide initially in solution. Fresh samples diluted fivefold in phosphate-buffered saline (PBS) and analyzed by means of DLS indicated particles of hydrodynamic diameter 132 ± 42 nm (Fig. 2C). The role of the disulfide linkage in the assemblies was confirmed by adding the disulfide-reducing agent tris(2-carboxyethyl)phosphine (TCEP) to the suspension. This ruptured the particles, producing smaller structures of diameter 2.3 ± 0.9 nm (Fig. 2C) similar to that observed for a mixture of CC-Tri3 and CC-Di-AB (2.5 ± 0.6 nm) (fig. S10).

Scanning electron microscopy (SEM) revealed closed spherical objects of similar diameter (97 ± 19 nm, n = 135 particles) (Fig. 3A and fig. S11). Although the particles appear as aggregates in these particular micrographs, they dispersed in solution (Fig. 2C) and separated when deposited on porous membranes (fig. S12). Tapping-mode atomic force microscopy (TM-AFM) was performed on particles deposited and dried onto mica. These particles were flattened disks 9.2 ± 1.0 nm thick (averaged from scans over five particles) with diameters of 95 ± 14 nm (from four measurements each on five particles) (Fig. 3, C and D, and fig. S13). As the coiled-coil modules are estimated to be ~3 nm in length, the observed thickness of these disks is strong evidence that in solution, the spheres are hollow and unilamellar rather than being solid, multiwalled, or onionlike structures. That is, they collapse upon drying, presumably releasing water through pores in the assembly. This fits our concept for the SAGEs: a

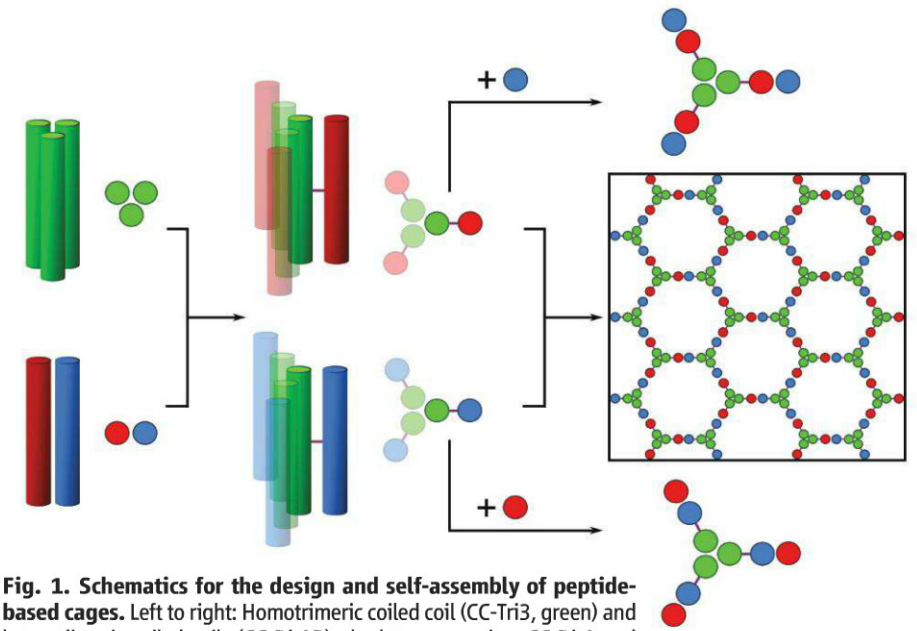


Fig. 1. Schematics for the design and self-assembly of peptide-based cages. Left to right: Homotrimeric coiled coil (CC-Tri3, green) and heterodimeric coiled coils (CC-Di-AB); the latter comprises CC-Di-A and CC-Di-B, colored red and blue, respectively. CC-Tri and CC-Di-AB are linked via asymmetric disulfide bonds (purple lines) to render hub A (green-red) and hub B (green-blue). Mixing hub A with CC-Di-B, or hub B with CC-Di-A produces discrete nine-helix assemblies, whereas mixing the hubs directly produces a hexagonal network, which should close.

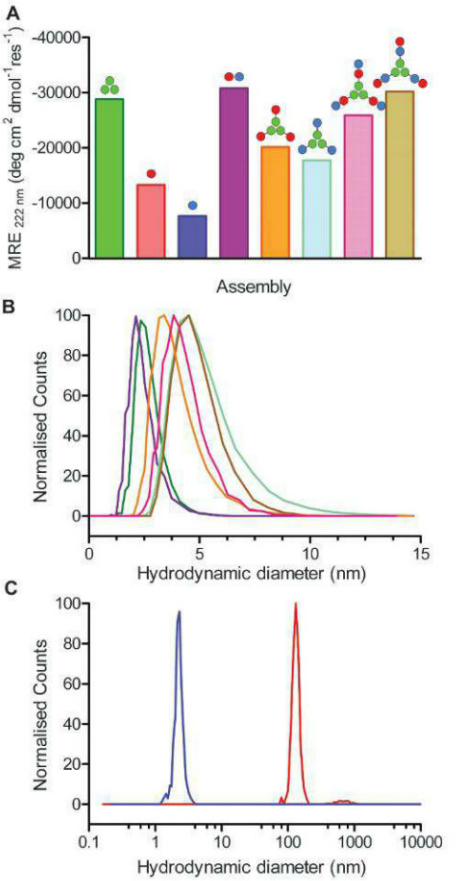


Fig. 2. Solution-phase characterization of the designed peptide modules, hubs, and discrete assemblies. (**A**) Mean residue ellipticity (MRE) at 222 nm (MRE_{222}) from CD spectroscopy in PBS at 20°C for CC-Tri3 (50 μM, green); CC-Di-A (50 μM, red); CC-Di-B (50 μM, blue); the mixture of CC-Di-A and CC-Di-B (CC-Di-AB) (50 μM + 50 μM, purple); hub A (CC-Tri3–CC-Di-A, 50 μM, orange); hub B (CC-Tri3–CC-Di-B, 50 μM, cyan); hub A plus CC-Di-B (50 μM + 50 μM, pink); hub B plus CC-Di-A (50 μM + 50 μM, brown). Full CD spectra and thermal denaturation curves are provided in fig. S3 (*20*). The leucine-zipper peptide, GCN4-p1, gives a benchmark MRE_{222} for 100% α-helix of ≈−36,000 deg cm^2 $dmol^{-1}$ res^{-1} (*25*). (**B**) Hydrodynamic diameters of the coiled-coil modules, hubs, and nine-chain terminated assemblies as determined with dynamic light scattering; colors are as per (A). Expanded DLS data are given in fig. S5. (**C**) DLS data for the assembled SAGE particles before (red) and after (blue) treatment with TCEP.

folded sheet comprising a hexagonal network of peptides (Fig. 1).

Lateral molecular-force microscopy (LMFM) with optical feedback was used in a noncontact regime to explore the assemblies in solution (*21*). Again, this showed approximately spherical objects (diameter 79 ± 12 nm (n = 19 measurements); height 82 ± 16 nm (averaged from scans over six particles) (Fig. 3E). These dimensions are similar to those found with SEM, which should be ~10 nm larger because of the sputtered metal coating estimated from the manufacturer's technical notes to be ~5 nm thick (*20*, *22*). Moreover, the LMFM revealed ultrastructure on the surfaces of the assemblies, notably clear hexagonal shapes (Fig. 4, G to I). The edges of the hexagons averaged 7 ± 2 nm (n = 22 edges), although such x and y dimensions in scanning probe microscopies are tip dependent and are not as reliable as measurements made in z.

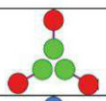
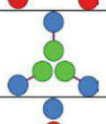
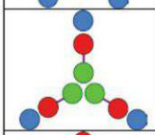
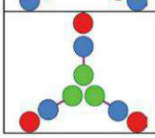
Our observations of closed spheres with a tight size distribution, confirmed by three independent methods, raises two immediate questions: How do the hexagonal networks fold and close, and why are the resulting closed structures so uniform in size?

The first question arises because rigid hexagonal networks should form flat assembles (like a graphite sheet), and closing a sphere (as illustrated by a football) cannot be achieved with hexagons alone and requires, for example, 12 pentagons. However, the coiled-coil modules and hubs of the SAGEs are more flexible, and the assemblies that they produce may tolerate imperfections required to close. Such imperfections, which are inevitable when closing such structures, could include a few mismatched hub pairings, rather than the perfect hexagonal array shown in Fig. 1.

Closing the particles may be driven by thermodynamic and geometric constraints. Regarding thermodynamics, the hubs are designed to associate with their complementary partners, which has two consequences: (i) hubs from solution coassemble to grow the network, and (ii) these expanding edges have unsatisfied coiled coils, which drive the sheets to close and satisfy as many coiled-coil interactions as possible. In terms of geometry, it is likely that there is some intrinsic tendency for the hubs to prefer tripod-like structures, with arms arranged at less than 120°, creating curvature. We tested these ideas computationally and experimentally as follows.

Complete SAGEs are too large for atomistic simulations, so we modeled smaller fragments of the hexagonal network. From x-ray crystal structures (*18*) and standard coiled-coil parameters (*23*), we generated an array of 19 tessellated hexagons built from CC-Tri3 and CC-Di-AB modules, and with 306 chains in total (Fig. 4, A and C). After 5 ns of molecular dynamics (MD) in water, uniform curvature was evident in both the x and y directions (Fig. 4, B and D, and movie S1). This was reproducible: In this, and multiple MD simulations for smaller seven-hexagon networks, the CC-Tri3 modules remained perpendicular to the curved surface with their N-termini always facing "out."

Table 1. Summary of solution-phase biophysical data. All biophysical data were obtained from samples of 50 μM concentration of each peptide component listed and in PBS at pH 7.4.

Assembly		MRE_{222}* (deg cm^2 $dmol^{-1}$ res^{-1})	T_M† (°C)	Mass of assembly by AUC (daltons) [95% CI] (calculated mass)	Hydrodynamic diameter by DLS (nm)
	CC-Tri3	-28,780	65	7944 [7858–8009] (7960)	2.6 ± 0.5
	CC-Di-A	-13,294	ND	ND	ND
	CC-Di-B	-7,647	ND	ND	ND
	CC-Di-AB	-30,746	51	5350 [5289–5442] (5,590)	2.3 ± 0.6
	*(hub A)* CC-Tri3—CC-Di-A	-20,168	68	16,910 [16720–17110] (16,238)	3.7 ± 0.6
	*(hub B)* CC-Tri3—CC-Di-B	-17,764	56	15,810 [15,662–15,963] (16,153)	5.0 ± 0.3
	*hub A* + CC-Di-B	-25,894	54	24,980 [24,795–25,143] (24,505)	4.1 ± 0.6
	*hub B* + CC-Di-A	-30,218	57	31,690 [31,574–31,793] (24,505)	4.7 ± 0.2

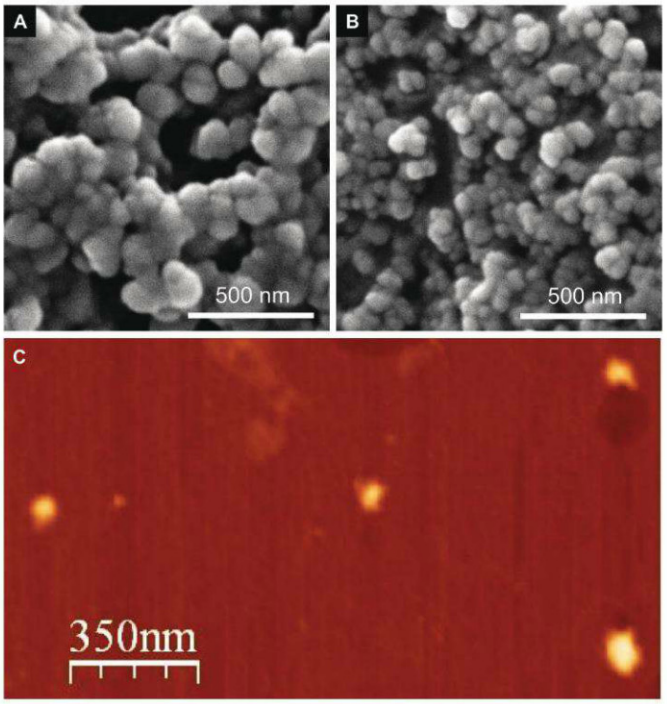
*MRE observed at 222 nm, 20°C. †Midpoint, or point of inflection, in sigmoidal thermal unfolding curves recorded as the change in MRE_{222} while ramping the temperature from 5 to 90°C at 40°C/hour.

A sphere of diameter 100 nm has a girth of ≈314 nm, corresponding to ≈40 equatorial hexagons. Thus, each hexagon is required to be wedge-shaped subtending an angle of ≈10° at the center of the sphere. Further examination of the MD trajectories and retrospective inspection of the designed sequences suggest a molecular interpretation for this wedging: The disulphide bridges linking the coiled coils are slightly offset toward the C-termini, and each peptide has a positively charged lysine residue at the f site between these bridges and the N-termini (fig. S1). As borne out by the MD, the positively charged lysine residues repel each other, whereas the disulfide bonds act as a tether. The overall effect is to splay the collective N-termini of each coiled-coil unit apart, resulting in wedge-shaped hubs (fig. S14) and producing local and then global curvature.

The question regarding the tight size distribution of the SAGEs is more difficult to rationalize, although this is likely to involve elements such as hub rigidity, the proportion of imperfections required to close a sphere, and entropic factors. To examine how hub rigidity and any preferred local curvature may vary, we analyzed multiple MD simulations of seven-hexagon tessellates from different starting conditions. After 10-ns simulations, the hub-hub angle approached equilibrium settling to 33.9 ± 17.2° (fig. S15). Clearly, the simulations overestimate the local, and therefore global, curvature. Nonetheless, the 10° angle estimated from the experiments is sampled in the simulations (fig. S15).

To exploit this apparent flexibility, and to test the importance of burying unsatisfied edges en route to closure, we attempted to engineer smaller SAGE particles. We prepared an additional heterodimer module, CC-Di-AIBI (table S1). In these peptides, Asn→Ile mutations were made at complementary ***a*** sites in the hydrophobic face so as to give a variant with more than two orders of magnitude higher affinity than the CC-Di-AB parent (*18*); otherwise, we do not expect this change to alter coiled-coil or hub structure or geometry. Thus, the free-energy penalty associated with unsatisfied edges, and proposed to drive closure, should be higher for the variant. When compared by means of SEM (Fig. 3, A and B), the parent SAGE particles had diameters of 97 ± 19 nm (n = 135 particles), whereas those incorporating the variant had diameters of 68 ± 12 nm (n = 97 particles) ($P < 0.001$). This translates to the latter having about half the surface area and provides strong evidence that satisfying coiled-coil interactions on the edge of a growing disk is a key driving force in closing assemblies. Moreover, it illustrates another advantage of our modular design strategy—namely, that altering the K_d of the individual coiled coils can be used to control SAGE size.

The SAGE concept, although inspired by natural examples, offers routes to closed systems of reduced complexity, with the potential for encapsulation. Because the components are modular, interchangeable, and bear termini and side



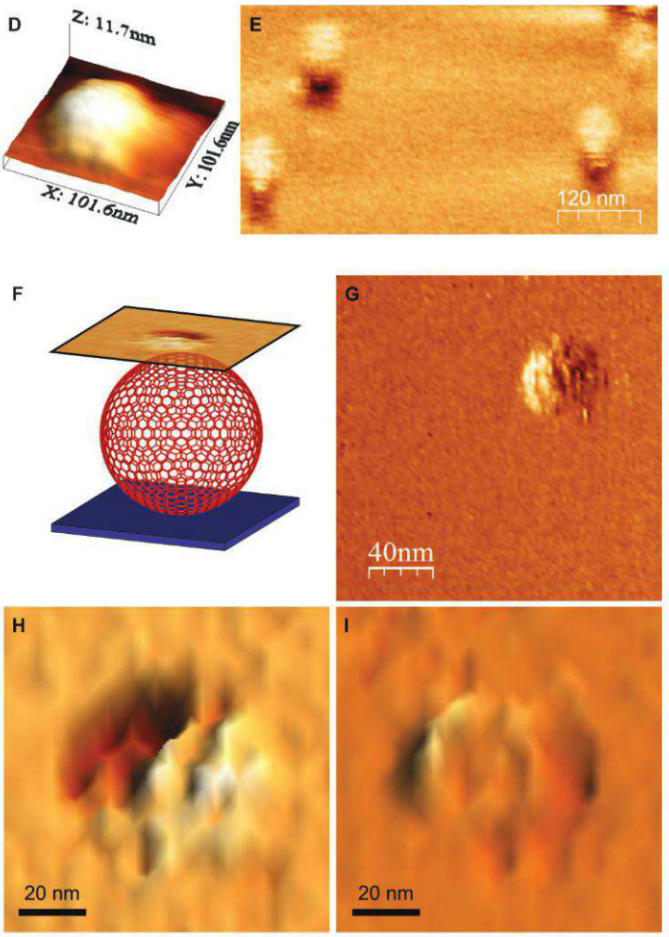
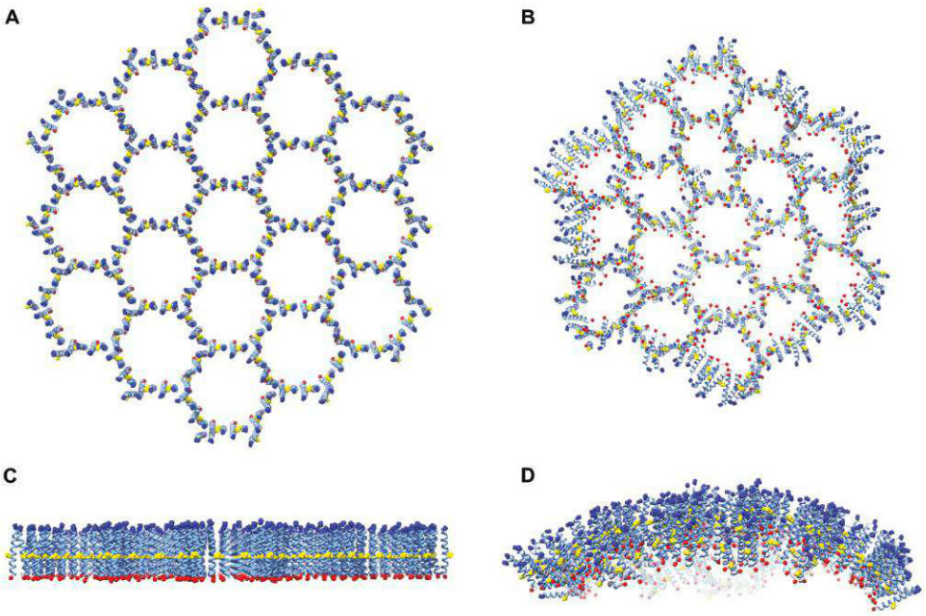


Fig. 3. Electron and force microscopy of SAGEs. (**A** and **B**) Scanning electron micrographs of the SAGE particles. (A) Images recorded after mixing hub A and hub B in PBS, pH 7.4 (to final concentrations of 50 μM in each of the component peptides, CC-Tri3–CC-Di-A and CC-Tri3–CC-Di-B; 16.7 μM of each hub). After 1 hour, resuspended material was transferred to a carbon-coated stub and sputter-coated with gold/palladium before imaging. (B) Smaller SAGE particles formed by mixing CC-Tri3–CC-Di-AI and CC-Tri3–CC-Di-BI. Full images are provided in fig. S11. (**C** to **I**) Scanning probe microscopy of the SAGEs. (C) TM-AFM scan of four individual collapsed SAGEs dried onto a mica substrate. (D) Three-dimensional representation of the topography measured over a single collapsed SAGE via TM-AFM. (E) LMFM scan recorded in liquid in a noncontact regime with a constant separation of the tip from the glass substrate. (F) Schematic representation of the LMFM scanning regime. An optical feedback maintained the vertically oriented cantilever at a constant separation from the substrate. Mapping the shear-force interaction nanometers above the SAGE allowed an image to be collected in a noncontact mode. [(G) to (I)] LMFM images of the hexagonal ultrastructure on the surfaces of hydrated SAGEs.

Fig. 4. Molecular dynamics of a model for the SAGEs. (**A** and **C**) Atomistic molecular model of 19 tessellated hexagons comprising hub A and hub B across three perspectives. These were generated from models for CC-Tri3 and CC-Di-AB aligned "back to back" along the three- and twofold symmetry axes and linked through disulfide bonds. As for the synthesized peptides, the termini of the modeled peptides were capped and are represented in blue (acylated *N*-termini) and red (amidated *C*-termini), with the Cys-Cys disulfide linkers colored yellow. (**B** and **D**) Curvature observed after 5-ns MD simulations. These were performed under periodic boundary conditions by using explicit water (TIP3P) at pH 7 and with 150 mM NaCl present (movie S1) (*20*).



chains that could be derivatized, it should be possible to tune their properties for applications such as vehicles for drug and biomolecule delivery, cages for trapping functional enzyme cascades that allow flux of starting materials and products, components of sensing systems, and new frameworks for the development of protocells (*24*).

References and Notes

1. H. F. Lodish, *Molecular Cell Biology* (W.H. Freeman, New York, ed. 6, 2008).
2. S. Tanaka *et al.*, *Science* **319**, 1083 (2008).
3. C. M. Agapakis, P. M. Boyle, P. A. Silver, *Nat. Chem. Biol.* **8**, 527 (2012).
4. D. A. Hammer, N. P. Kamat, *FEBS Lett.* **586**, 2882 (2012).
5. M. Uchida *et al.*, *Adv. Mater.* **19**, 1025 (2007).
6. R. M. Capito, H. S. Azevedo, Y. S. Velichko, A. Mata, S. I. Stupp, *Science* **319**, 1812 (2008).
7. F. Boato *et al.*, *Angew. Chem. Int. Ed. Engl.* **46**, 9015 (2007).
8. S. Raman, G. Machaidze, A. Lustig, U. Aebi, P. Burkhard, *Nanomedicine* **2**, 95 (2006).
9. N. P. King *et al.*, *Science* **336**, 1171 (2012).
10. Y. T. Lai, D. Cascio, T. O. Yeates, *Science* **336**, 1129 (2012).
11. B. Wörsdörfer, K. J. Woycechowsky, D. Hilvert, *Science* **331**, 589 (2011).
12. J. C. Sinclair, K. M. Davies, C. Vénien-Bryan, M. E. M. Noble, *Nat. Nanotechnol.* **6**, 558 (2011).
13. J. D. Brodin *et al.*, *Nat. Chem.* **4**, 375 (2012).
14. M. M. Pires, J. Lee, D. Ernenwein, J. Chmielewski, *Langmuir* **28**, 1993 (2012).
15. E. H. C. Bromley, K. Channon, E. Moutevelis, D. N. Woolfson, *ACS Chem. Biol.* **3**, 38 (2008).
16. A. N. Lupas, M. Gruber, *Adv. Protein Chem.* **70**, 37 (2005).
17. D. N. Woolfson, *Adv. Protein Chem.* **70**, 79 (2005).
18. J. M. Fletcher *et al.*, *ACS Synth. Biol.* **1**, 240 (2012).
19. E. H. C. Bromley, R. B. Sessions, A. R. Thomson, D. N. Woolfson, *J. Am. Chem. Soc.* **131**, 928 (2009).
20. Materials and methods are available as supplementary materials on *Science* Online.
21. R. L. Harniman *et al.*, *Nanotechnology* **23**, 085703 (2012).
22. www.quorumtech.com/pdf/SputterCoatingTechniques/Sputter_coating_technical_brief.pdf
23. G. Offer, R. Sessions, *J. Mol. Biol.* **249**, 967 (1995).
24. S. S. Mansy *et al.*, *Nature* **454**, 122 (2008).
25. E. K. O'Shea, R. Rutkowski, P. S. Kim, *Science* **243**, 538 (1989).

Acknowledgments: We thank the Woolfson group for valuable discussions; the BBSRC for funding to D.N.W. and P.J.B. (BB/G008833/1) and for a studentship to A.L.B.; and the EPSRC for studentships to R.L.H. and T.H.S. J.M. is supported by a Wellcome University Award. We are grateful to the University of Bristol Advanced Computing Research Centre and the eInfraStructureSouth Consortium for high-performance computing; and Jonathan Jones and the EM Unit (EMU) School of Chemistry University of Bristol for EM access. The peptides described have been added to the Pcomp database for synthetic biology (http://coiledcoils.chm.bris.ac.uk/pcomp/index.php).

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1233936/DC1
Materials and Methods
Figs. S1 to S15
Tables S1 to S3
References (*26–38*)
Movie S1

11 December 2012; accepted 11 March 2013
Published online 11 April 2013;
10.1126/science.1233936

Amplifying Genetic Logic Gates

Jerome Bonnet, Peter Yin,* Monica E. Ortiz, Pakpoom Subsoontorn, Drew Endy†

Organisms must process information encoded via developmental and environmental signals to survive and reproduce. Researchers have also engineered synthetic genetic logic to realize simpler, independent control of biological processes. We developed a three-terminal device architecture, termed the transcriptor, that uses bacteriophage serine integrases to control the flow of RNA polymerase along DNA. Integrase-mediated inversion or deletion of DNA encoding transcription terminators or a promoter modulates transcription rates. We realized permanent amplifying AND, NAND, OR, XOR, NOR, and XNOR gates actuated across common control signal ranges and sequential logic supporting autonomous cell-cell communication of DNA encoding distinct logic-gate states. The single-layer digital logic architecture developed here enables engineering of amplifying logic gates to control transcription rates within and across diverse organisms.

Researchers have used genetically encoded logic, data storage, and cell-cell communication to study and reprogram living systems, explore biomolecular computing, and improve cellular therapeutics (*1–9*). Most approaches to engineering cell-based logic champion two-terminal device architectures upon which gate-gate layering, similar to conventional electronics, is used to realize all logic functions (*10*, *11*). Despite recent advances (*11*, *12*), such designs are difficult to scale because of problems associated with reusing regulatory molecules within the self-mixing environments of individual cells. As representative examples, a single-cell two-input "exclusive or" (XOR) gate, a function whose output is high only if the inputs are different, required controlled expression of four gate-specific regulatory molecules from four plasmids (*12*); an amplifying "exclusive nor" (XNOR) gate, high output only if inputs are equal, has not been demonstrated within single cells (*10*).

Department of Bioengineering, Y2E2-269B, 473 Via Ortega, Stanford, CA 94305–4201, USA.

*Present address: Department of Biology, University of Pennsylvania, Philadelphia, PA 19104, USA.
†Corresponding author. E-mail: endy@stanford.edu

We instead sought a device architecture in which the same regulatory molecules could be simply reused to implement all logic gates within a single logic layer (*13*). We also sought to decouple the signals controlling gate switching from gate inputs and outputs. Realizing both goals would enable straightforward engineering of distinct gates with constant switching thresholds and support signal gain and amplification if desired. Lastly, we wanted all gate signals to be encoded via a common signal carrier supporting connectivity within natural systems and across a diverse family of engineered genetic devices (*14*).

We combined earlier concepts (*14–17*) to invent a transistor-like three terminal device (*18*) termed the transcriptor. Independent control signals govern transcriptor logic elements that regulate transcriptional "current," defined by the flow of RNA polymerase along DNA (Fig. 1A). Gate input and output signals are transcription rates at positions on DNA marking logic element boundaries. Logic elements use asymmetric transcription terminators as reversible check valves that disrupt RNA polymerase flow in only one of two possible orientations (Fig. 1B). Recombinases catalyze unidirectional inversion of DNA within opposing recognition sites (Fig. 1B) or deletion of DNA between aligned sites (Fig. 1C), providing independent control over the orientation or presence of one or more terminators. Stated differently, we developed a device architecture similar to a transistor but leveraged unique properties of genetic regulation to implement all gates without requiring that multiple instances of simpler gates be connected in series (i.e., without layering) (*18*, *19*).

For example, a transcriptor XOR logic element requires bracketing one asymmetric transcription terminator with two pairs of opposing recombination sites recognized by independent integrases (Fig. 1D). If neither integrase is expressed, then the terminator blocks transcription (Fig. 1D, top). Expression of either integrase alone inverts the DNA encoding the terminator and allows transcription to flow through the transcriptor (Fig. 1D, middle). Expression of both integrases inverts and then restores the original orientation of the terminator, again blocking transcription (Fig. 1D, bottom). A complete XOR gate requires placing an XOR logic element within a three-terminal device in which integrase expression is controlled by two independent control signals (Fig. 1E).

We designed additional transcriptor logic elements encoding Boolean OR, NOR, XNOR, and AND functions for use within a common gate architecture (Fig. 2 and fig. S1). Straightforward changes only to logic element DNA were sufficient to design functionally distinct gates expected to be responsive to identical control signals. Designing a transcriptor-only "not and" (NAND) element, low output only if both control signals are high, proved more challenging. We instead used a hybrid architecture that combines flipping of a terminator along with a constitutive promoter. Although noncanonical, the NAND gate still responds to the same control signals while exhibiting varied output levels (below).

Transcriptor elements allowing unidirectional processing of DNA produce permanent gates that implement write-once logic operations. However, recombination directionality factors (RDFs) can reverse DNA inversion by integrases (*20*). We designed rewritable transcriptor elements in which controlled expression of RDFs, given constitutive integrase expression, should implement reversible logic and demonstrated a reversible buffer gate controlled by a single RDF (figs. S2 and S3); reversible gates require the same number of regulated factors as permanent gates (e.g., two RDFs versus two integrases). Permanent gates are useful for applications requiring tracking or processing of historical signals (e.g., terminal differentiation during development or accumulated responses to asynchronous environmental cues), whereas reversible gates support multicycle computing (e.g., a synchronized cell division cycle counter).

We selected unidirectional serine integrases from bacteriophages TP901-1 and Bxb1 to control gate switching; these recombinases do not require host cofactors and have been shown to function in bacteria, fungi, plants, and animals (*21*). We recently implemented a rewritable digital latch by using the Bxb1 integrase and RDF to repeatedly flip a DNA memory element between two states (*5*); this class of latches are controlled by continuous transcription signals (i.e., analog inputs) that produce recombinase proteins sufficient to flip (or not) a DNA element (i.e., digital output) and thus can be abstracted and reused (*22*) as analog-to-digital converters.

We made a recombination control plasmid expressing the TP901-1 and Bxb1 integrases under the control of exogenous arabinose (ara) and anhydrotetracycline (aTc) induction, respectively (*23*). We measured the propensities of TP901-1 and Bxb1 integrases to recognize and process specific DNA recombination sites (fig. S4A). We fitted abstracted Hill functions representing observed DNA processing propensities given increasing expression levels for each integrase alone (fig. S4B). We defined logic function models specific to each gate (fig. S4C) and predicted the expected behavior of multi-input gates directly from single-integrase Hill functions (Fig. 2).

We constructed low-copy plasmids encoding AND, OR, XOR, NAND, NOR, and XNOR logic elements between a standard strong prokaryotic promoter (input signal source) and green fluorescent protein (GFP) expression cassette (indirect output signal reporter) (*23*). We measured bulk GFP levels from bacterial cultures expressing varying amounts of TP901-1 and Bxb1 integrases controlling the six Boolean gates. Observed GFP expression patterns were well matched to predictions for all gates (Fig. 2); exact output levels varied among and within some gates (described below).

Transcriptor-based gates use discrete enzymatic processing of DNA to modulate RNA polymerase flow through logic elements. We thus expected that single cells might exhibit discontinuous (e.g., all or none) responses to small changes in control signals. For example, we measured fluorescence output distributions among single cells exposed to low or high control signals and found a single threshold defining distinct low/high outputs across all gates (Fig. 2, red vertical line). To better study control signal digitization (i.e., the extent to which gate outputs are more digital than gate control signals across small changes in control signals), we compared changes in gate outputs to increasing control signals by using a common reporter (figs. S5 to S8). For example, we found that XOR gates switch completely between 0.2 and 2 ng/ml aTc and 0.0001 to 0.001% ara, whereas both control signals increase gradually across these inducer ranges (Fig. 3, A and B, and figs. S6 to S8). We defined a digitization error rate as the combined probability of scoring false high or low gate outputs in response to intermediate control signal changes, optimized thresholds for controllers and gates that best discriminate between putative low and high outputs, and quantified digitization error rates for each gate. AND, OR, XOR, and XNOR gates digitized aTc-induced control signals to varying degrees, whereas AND, OR, NOR, XOR, and XNOR gates digitized ara-induced signals (Fig. 3, C and D); no gates reduced digitization (Fig. 3, C and D), and all gates realized digital outputs in response to low/high control signals (Fig. 2).

Changes in gate outputs must be compared directly to changes in gate control signals to determine whether gates function as amplifiers (*24*, *25*). We calculated population-average GFP levels for each control signal and gate outputs (fig. S5). We directly compared changes in gate outputs to the changes in gate control signals needed to activate gate switching, both for absolute (figs. S9 and S10) and normalized (Fig. 4) expression levels. We evaluated all gates across control-signal ranges needed to drive the least-

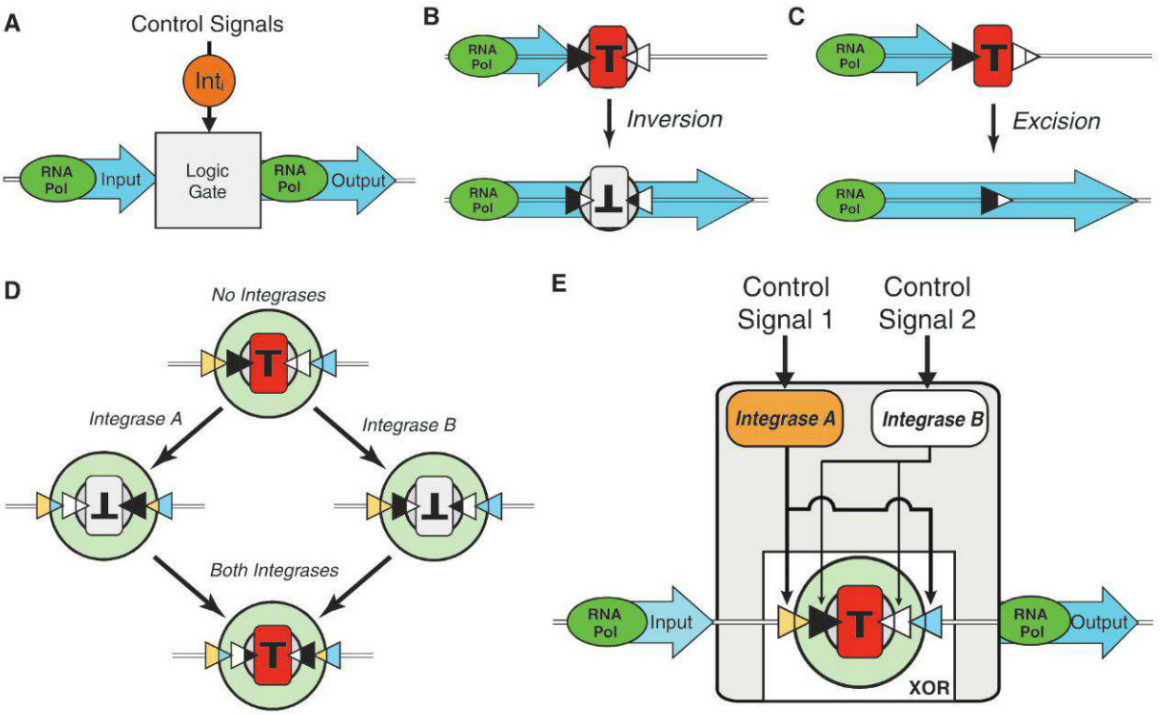


Fig. 1. Using transcriptors to implement three-terminal Boolean integrase logic gates. (A) Three-terminal transcriptor-based gates use integrase (Int) control signals to modulate RNA polymerase (RNA Pol) flow between a separate gate input and output. **(B)** A canonical transcriptor element wherein an asymmetric transcription terminator (T) blocks transcriptional current in one orientation (red) or, when flipped, the opposite orientation (gray). Opposing recombination sites (black/white triangles) flank the terminator and direct flipping by an integrase. **(C)** Integrases can also excise DNA between aligned recombination sites. **(D)** State diagram for a transcriptor "exclusive or" (XOR) logic element: Opposing sites recognized by two independent integrases (blue/orange and black/white) are nested and flank one terminator. Recombination can produce four distinct states controlling terminator orientation. **(E)** The logic element from (D) within a three-terminal Boolean integrase XOR gate such that gate output is high only if control signals are different.

responsive gate (NAND) and did not normalize outputs via subtraction of lowest values, which otherwise greatly increases fold-change estimates. All gates generate increased absolute and fold-change differences in expressed output protein levels relative to those produced from the integrase controllers (Fig. 4 and figs. S9 and S10).

We confirmed that permanent transcriptor-based gates support sequential logic based on the heritable storage of logic element states in response to asynchronous control signals. For example, cells encoding AND and XNOR gates were exposed to various patterns of integrase control signals, recording and generating appropriate outputs across ~40 cell doublings (fig. S11). Building from these results, we engineered cell-cell communication of DNA (*7*) encoding logic gates at different stages of gate activation (fig. S12), a feature unique to gates whose operation involves direct processing of DNA. We also assayed recombination response times, finding that 15-min control-signal pulses were sufficient to activate integrase-mediated switching (fig. S13 and movie S1).

During the course of responding to reviewer questions and preparing the final version of this manuscript, a system of two terminal logic gates was described by Siuti *et al*. based on flipping terminator, promoter, and gene elements (*26*). The family of Boolean integrase logic gates introduced here differs in the consistent use of a three-terminal device architecture that decouples logic-gate operation from both input and output signals, enabling simple tuning via changes to the transcription input signal (fig. S14) and ready reuse (e.g., reprogramming of natural transcription

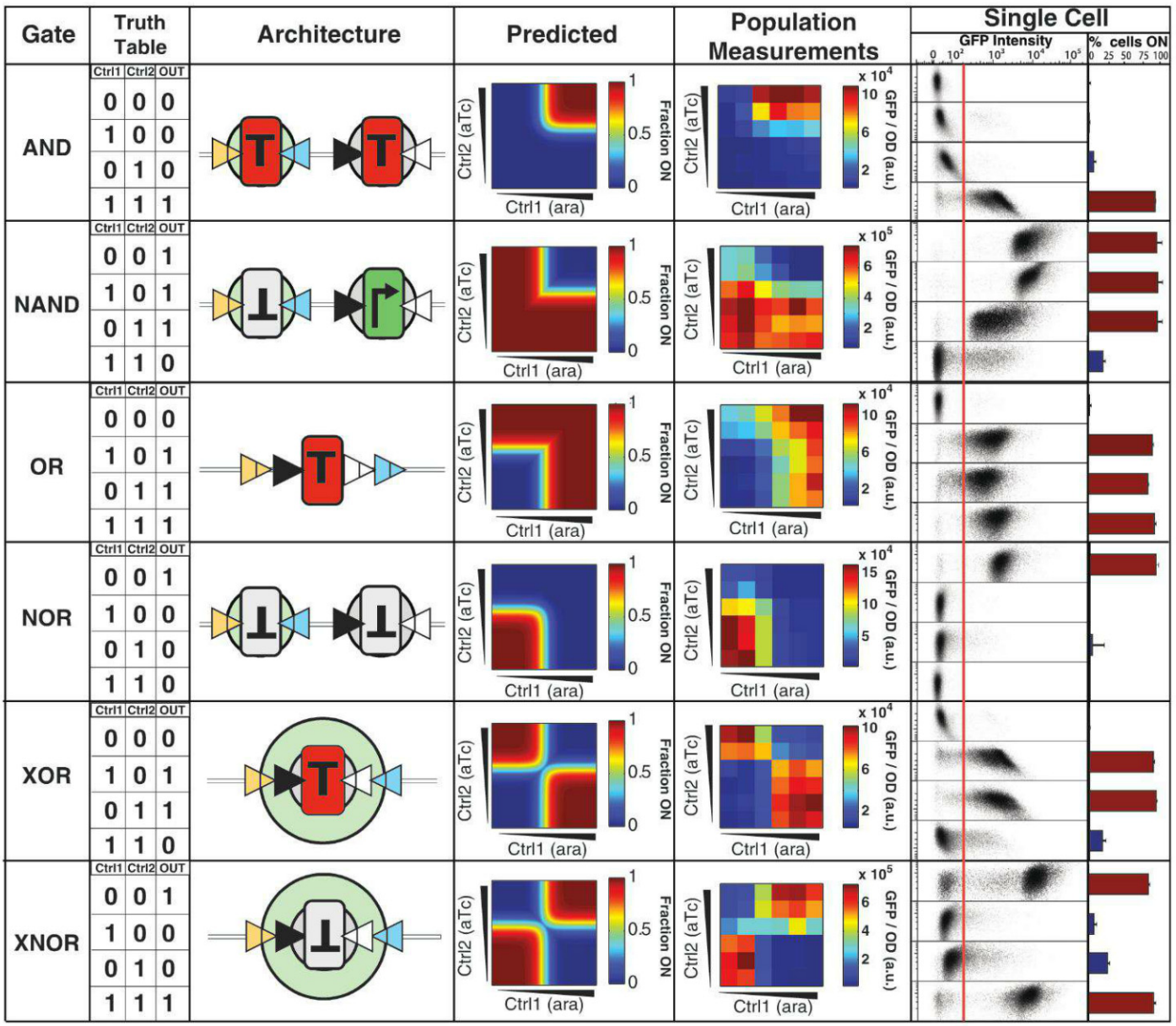


Fig. 2. Predicted and observed logic-gate performance plus digital-output thresholding. (Gate) Boolean logic functions. (Truth Table) Logical relationships between control (Ctrl) signals and output. (Architecture) Transcriptor element configurations for use in permanent Boolean integrase logic gates, including a constitutive promoter (NAND, green rightward arrow). (Predicted) Expected fraction of cells containing gates in a high-output state (heat map) calculated as described (*23*). (Population Measurements) Gate outputs assayed via plate reader for bulk cultures (*23*). Inducer concentrations are 0, 0.02, 0.2, 2, 20, and 200 ng/ml for aTc and 0, 0.0001, 0.001, 0.01, 0.1, and 1% weight/volume (w/v) for ara. (Single Cell, GFP Intensity) Distribution of gate outputs [*x* axis GFP output measured in arbitrary units (a.u.); *y* axis side scatter] among single cells responding to control signals, as per gate-specific truth tables. A common output threshold segregates low/high outputs across all gates (vertical red line). Inducer concentrations are 200 ng/ml for aTc and 1% w/v for ara. (% Cells On) Fraction of cells encoding high outputs when scored by using a common output threshold.

Fig. 3. Digitization of control signals. (**A**) Distribution of XOR outputs among single cells (red contours) responding to an increasing control signal (blue contours). Each contour interval encompasses 5% of all cells; thick contours surround 50% of the total population. *y* axis, GFP output measured in arbitrary units (a.u.); *x* axis, side scatter at noted inducer concentrations. (**B**) As in (A) but for the second control signal. (**C**) Gate switching and digitization errors across an intermediate control signal change (0.2 to 2 ng/ml aTc, left/right of each frame). Gate-specific digitization thresholds (horizontal bars) were optimized and used to quantify fractions of false high cells given a low control signal and vice versa. Numbers within frames are high-given-low and low-given-high error rates. Numbers above boxes are combined error rates with standard deviation from three independent experiments. (**D**) As in (C), but in response to an intermediate change in the second control signal (1×10^{-4} to 0.5×10^{-2} ara, left/right of each frame).

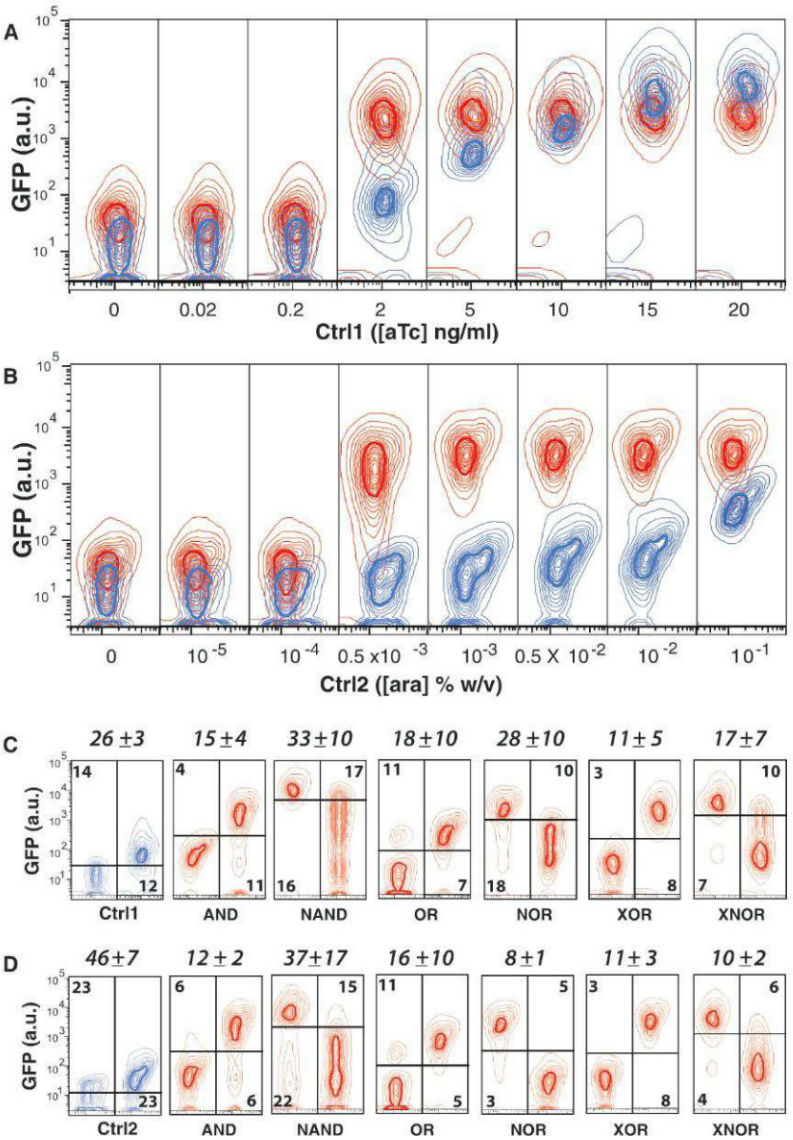
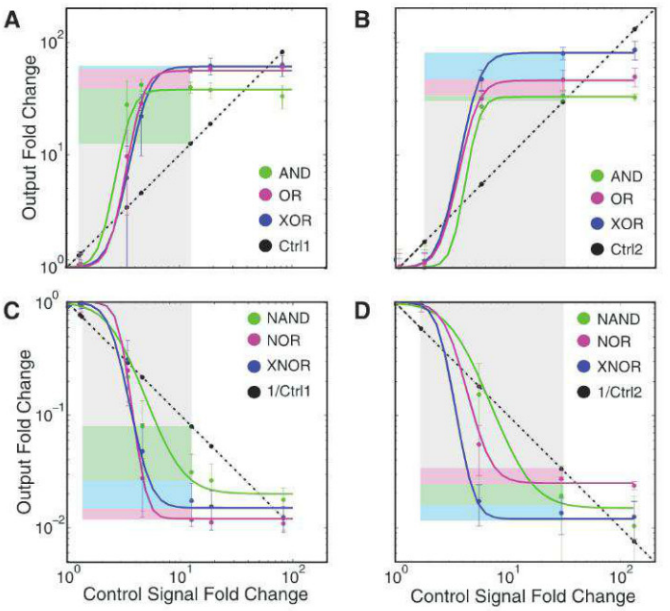


Fig. 4. Gain and amplification across common control signal ranges. Population-average response of amplifier gates to (**A**) increasing ara-mediated expression of TP901-1 integrase and (**B**) increasing aTc-mediated expression of Bxb1 integrase. Changes in output GFP levels produced by gates (colored lines) are directly compared to changes in control signals required for gate switching (increasing straight dashed lines). The response of each control signal to itself (gray boxes) is shown to highlight gate-specific amplification of control signals (colored boxes). (**C** and **D**) Responses of inverting amplifier gates. As in (A) and (B), except that fold changes for control signals are inverted (decreasing straight dashed lines) (figs. S8 and S10). Error bars indicate SD across three independent experiments.



systems by seamless integration of transcriptor logic elements within natural operons). Also, by separating gate inputs from gate control signals and by using a strong input signal modulated by an efficient asymmetric terminator, we were able to demonstrate and quantify signal amplification for all gates (Figs. 3 and 4).

Output signal levels vary within and among the gates reported here (Figs. 2 and 3 and figs. S9 and S10), although not more so than existing genetic logic. We believe that most variation arises from differences in RNA secondary structures well known to influence mRNA stability and translation initiation rates (fig. S15); such variation might be eliminated by using recently reported mRNA processing methods (*24*, *27*). Further work is also required to realize precise level matching across all gates, and directed evolution of increasingly asymmetric terminators may be needed to reduce low output levels for most gates (fig. S10); additional gate-specific tuning of NAND would be required given its noncanonical logic element. Nevertheless, existing gates already support single-layer programmable digital logic, control-signal amplification, sequential logic, and cell-cell communication of intermediate logic states. Multi-input gates supporting high "fan-in" could be realized by using additional integrases (*28*) (fig. S16). Transcriptor-based gates can also likely be directly combined with other logic families to expand the power of engineered genetic computers. All logic gates and uses thereof demonstrated or disclosed here have been contributed to the public domain via the BioBrick Public Agreement (*29*).

References and Notes

1. B. Wang, M. Buck, *Trends Microbiol.* **20**, 376 (2012).
2. Y. Benenson, *Nat. Rev. Genet.* **13**, 455 (2012).
3. T. Miyamoto, S. Razavi, R. DeRose, T. Inoue, *ACS Synth. Biol.* **2**, 72 (2013).
4. D. R. Burrill, P. A. Silver, *Cell* **140**, 13 (2010).
5. J. Bonnet, P. Subsoontorn, D. Endy, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 8884 (2012).
6. S. Basu, Y. Gerchman, C. H. Collins, F. H. Arnold, R. Weiss, *Nature* **434**, 1130 (2005).
7. M. E. Ortiz, D. Endy, *J. Biol. Eng.* **6**, 16 (2012).
8. Y. Y. Chen, M. C. Jensen, C. D. Smolke, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 8531 (2010).
9. Z. Xie, L. Wroblewska, L. Prochazka, R. Weiss, Y. Benenson, *Science* **333**, 1307 (2011).
10. A. Tamsir, J. J. Tabor, C. A. Voigt, *Nature* **469**, 212 (2011).
11. T. S. Moon, C. Lou, A. Tamsir, B. C. Stanton, C. A. Voigt, *Nature* **491**, 249 (2012).
12. S. Ausländer, D. Ausländer, M. Müller, M. Wieland, M. Fussenegger, *Nature* **487**, 123 (2012).
13. For example, converting a NOR gate repressed by transcription factors to an OR gate activated by transcription factors requires changing how proteins interact with RNA polymerase (from competitive binding and occlusion to recruitment and initiation) and simultaneous reworking of the basal activity for core promoter elements (from a constitutively active promoter that can be repressed to a weak promoter that does not spontaneously initiate transcription yet that transcription factors activate).
14. C. Wadey, I. Deese, D. Endy, "Common signal carriers," in *Adventures in Synthetic Biology* (OpenWetWare and Nature Publishing Group New York, 2005), chap. 3; available online at http://hdl.handle.net/1721.1/46337.
15. T. S. Ham, S. K. Lee, J. D. Keasling, A. P. Arkin, *PLoS ONE* **3**, e2815 (2008).
16. A. E. Friedland *et al.*, *Science* **324**, 1199 (2009).
17. P. A. Varadarajan, D. Del Vecchio, *IEEE Trans. Nanobioscience* **8**, 281 (2009).
18. J. Bardeen, W. Brattain, *Phys. Rev.* **74**, 230 (1948).
19. With transistor-based logic, gates use a base, emitter, and collector architecture that classically only allows for control of electrical current at one point on a wire by a single signal. Transcriptor-based logic allows RNA polymerase flow at a single point on DNA to be controlled, in theory, by as many independent recombinases as needed.
20. J. A. Lewis, G. F. Hatfull, *Nucleic Acids Res.* **29**, 2205 (2001).
21. W. R. A. Brown, N. C. O. Lee, Z. Xu, M. C. M. Smith, *Methods* **53**, 372 (2011).
22. B. Canton, A. Labno, D. Endy, *Nat. Biotechnol.* **26**, 787 (2008).
23. Materials and methods are available as supplementary materials on *Science* Online.
24. C. Lou, B. Stanton, Y.-J. Chen, B. Munsky, C. A. Voigt, *Nat. Biotechnol.* **30**, 1137 (2012).
25. L. Pasotti, N. Politi, S. Zucca, M. G. Cusella De Angelis, P. Magni, *PLoS ONE* **7**, e39407 (2012).
26. P. Siuti, J. Yazbek, T. K. Lu, *Nat. Biotechnol.* (2013).
27. L. Qi, R. E. Haurwitz, W. Shao, J. A. Doudna, A. P. Arkin, *Nat. Biotechnol.* **30**, 1002 (2012).
28. G. F. Hatfull *et al.*, *J. Virol.* **86**, 2382 (2012).
29. https://biobricks.org/bpa/

Acknowledgments: We thank M. Juul, T. Knight, S. Kushner, C. Smolke, B. Townshend, the Endy and Smolke labs, and the Stanford Shared FACS Facility. Funding was provided by the NSF Synthetic Biology Engineering Research Center, Stanford Center for Longevity, Stanford Bio-X, the Townshend/Lamarre Family Foundation, and the Siebel Foundation. DNA sequences are available in GenBank (accession nos. KC529324 to KC529332). DNA constructs will be made available via Addgene.

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1232758/DC1
Materials and Methods
Figs. S1 to S15
Appendices S1 to S4
References (*30*–*44*)
Movie S1

14 November 2012; accepted 13 March 2013
Published online 28 March 2013;
10.1126/science.1232758

Controlled Flight of a Biologically Inspired, Insect-Scale Robot

Kevin Y. Ma,*† Pakpong Chirarattananon,† Sawyer B. Fuller, Robert J. Wood

Flies are among the most agile flying creatures on Earth. To mimic this aerial prowess in a similarly sized robot requires tiny, high-efficiency mechanical components that pose miniaturization challenges governed by force-scaling laws, suggesting unconventional solutions for propulsion, actuation, and manufacturing. To this end, we developed high-power-density piezoelectric flight muscles and a manufacturing methodology capable of rapidly prototyping articulated, flexure-based sub-millimeter mechanisms. We built an 80-milligram, insect-scale, flapping-wing robot modeled loosely on the morphology of flies. Using a modular approach to flight control that relies on limited information about the robot's dynamics, we demonstrated tethered but unconstrained stable hovering and basic controlled flight maneuvers. The result validates a sufficient suite of innovations for achieving artificial, insect-like flight.

Using flapping wings and tiny nervous systems, flying insects are able to perform sophisticated aerodynamic feats such as deftly avoiding a striking hand or landing on flowers buffeted by wind. How they perform these feats—from sensorimotor transduction to the unsteady aerodynamics of their wing motions—is just beginning to be understood (*1*–*3*), aided in part by simulation (*4*) and scaled models (*5*). Motivated by a desire for tiny flying robots with comparable maneuverability, we seek to create a robotic vehicle that mirrors these basic flight mechanics of flies. At the scale of flies, no such vehicle has been demonstrated to date because of the severe miniaturization challenges that must be overcome for an insect-sized device (*6*). Conventional technologies for macroscale aircraft propulsion and manufacturing are not viable for millimeter-scale robots because of inefficiencies that arise from force scaling, suggesting a biologically inspired solution based on flapping wings (*7*–*9*). Here, we report an aggregation of innovations in design, manufacturing, actuation, and control to create an insect-scale flying robot—a robotic fly—that successfully demonstrates tethered but unconstrained flight behavior reminiscent of flying insects.

For inspiration of form and function, we used *Diptera* (flies) as a model system because of the relative simplicity of the flight apparatus—flies by classification have only two wings—and the exemplary aerial agility that they exhibit. Dipteran flight has been well-studied (*5*, *10*–*18*), and it is understood that insect wings undergo a complex trajectory defined by three rotational degrees of freedom (*10*). This has been simplified in the robotic fly to a reciprocating flapping motion in which the wings' pitch rotation is regulated with passive compliant flexures (*19*)—an enabling simplification for mechanism design and manufacture. Key aspects of the oscillatory wing motion are the flapping frequency and wing stroke amplitude; the robotic fly achieves 120 Hz and 110°, respectively, similar to the 130-Hz wing beat

School of Engineering and Applied Sciences and the Wyss Institute for Biologically Inspired Engineering, Harvard University, Cambridge, MA 02138, USA.

*Corresponding author. E-mail: kevinma@seas.harvard.edu
†These authors contributed equally to this work.

frequency and 120°- to 150°-stroke amplitude in flies of comparable mass (*1*, *11*). The wing planform is inspired by the aspect ratio and area distribution of the hoverfly *Eristalis*, with wing area chosen to be consistent with similarly sized insects in terms of wing loading and peak wing velocities (*1*).

The mechanical components for a robotic fly require feature sizes between micrometers and centimeters—too large for silicon-based microelectromechanical systems (MEMS) and too small for conventional machining and assembly methods (*9*). Decreased feature size brings an increased dominance of surface forces, causing revolute joints or sliding surfaces to become inefficient or infeasible (*8*, *9*). Additionally, MEMS techniques, although sufficiently precise, are time-consuming, constrain material choice, and limit attainable geometries.

We developed a design and manufacturing methodology, "smart composite microstructures" (SCM), to address this void in mesoscale manufacturing (*9*). In SCM, material layers are bulk laser-micromachined and laminated together by using adhesives in a monolithic, planar fashion. Bulk micromachining results in feature sizes as small as 5 μm, and lamination allows virtually any material combination. Furthermore, the monolithic, parallel nature of this method facilitates mass production (Fig. 1A) and precision assembly through folding (*20*). All electromechanical elements of the robotic fly—including flight muscles, thorax, skeleton, and wings—were manufactured by use of SCM. Structural elements were created by using high stiffness-to-weight-ratio carbon fiber-reinforced composites, and articulation was achieved with polyimide film flexure hinges emulating low-friction revolute joints.

Flapping-wing flight is energetically costly (*1*), making power density and transduction efficiency vital metrics for flight muscle performance (*21*). Unfavorable scaling of magnetic forces limits the use of rotary electromagnetic motors (*8*, *21*), which are otherwise ubiquitous in larger robots. A survey of actuation technologies identified induced-strain materials, particularly piezoelectric ceramics, as most promising for oscillatory power delivery in insect-scale robots (*21*). Thus, for flight muscles we used voltage-driven piezoelectric bimorphs that can generate bidirectional forces, are compatible with our manufacturing methods, and are geometrically optimized for energy density (*22*).

The wing-flapping motion of the robotic fly is generated by a four-bar linkage that acts as a lever arm to amplify the small displacement of the piezoelectric flight muscle. A second degree of freedom is wing pitch rotation: This is relegated to a passive, elastic flexure hinge at the base of the wing. Inertial, aerodynamic, and elastic forces determine wing rotation as it interacts with the air (Fig. 1F) (*19*, *23*). This passive wing rotation mimics that observed in insects (*12*), although *Diptera* are known to possess additional musculature for active fine-tuning of rotation dynamics (*1*). The flapping motion, along with the passive pitch rotation of the wings, generates a downward propulsive force when averaged over a full stroke cycle. Thus, even with our design simplifications we can create wing kinematics that resemble wing motions in insect flight and generate sufficient lift forces for flight (*5*, *20*). Baseline wing kinematics are driven by sinusoidally exciting the piezoelectric flight muscle near the resonant frequency of the coupled muscle-thorax-wing system so as to minimize the energy expenditure on reactive power. Elastic energy storage in the robotic fly's flight apparatus parallels the energy storage observed in the thoracic mechanics of flies (*1*, *13*). Thrust modulation in the robot is achieved through amplitude modulation rather than frequency modulation so that the system

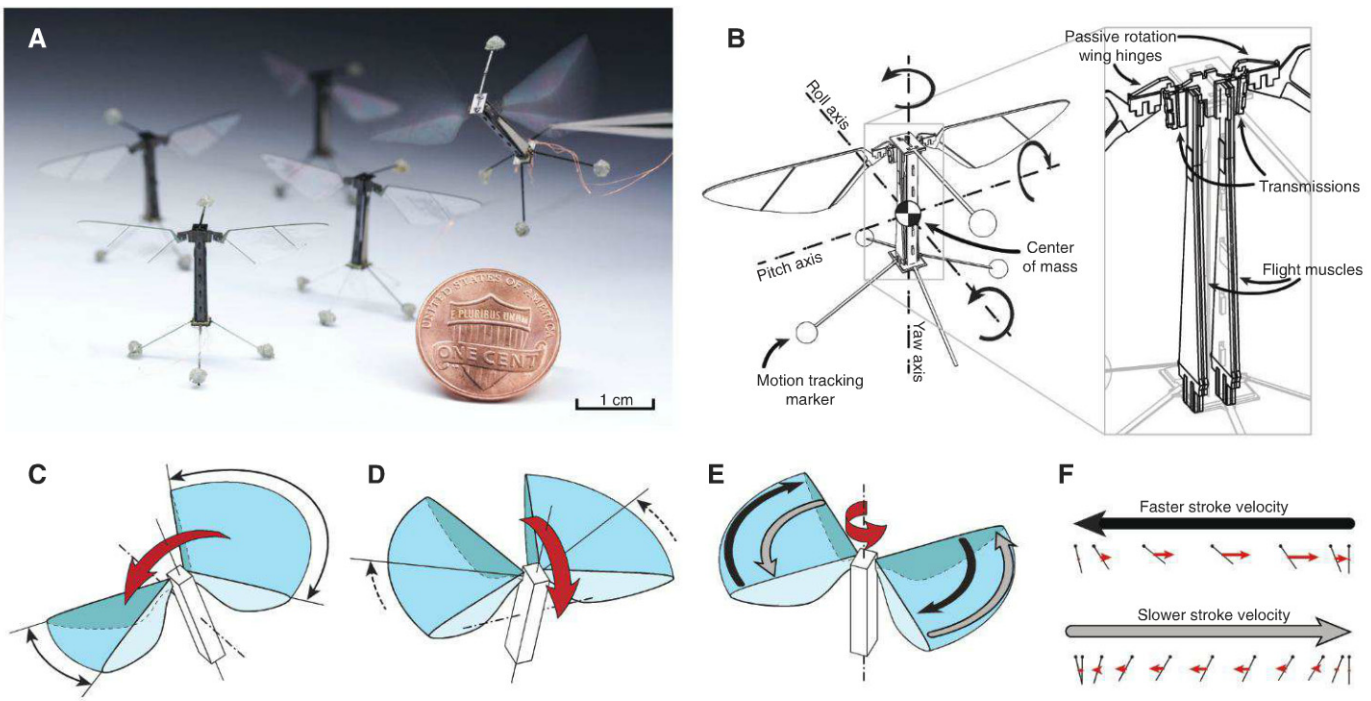


Fig. 1. A robotic fly design with a pair of independently actuated wings enables precise control of torques about three orthogonal axes. (**A**) Five individual robotic flies of identical design are shown alongside a U.S. penny for scale, demonstrating that the manufacturing process facilitates repeatability and mass production. (**B**) Body axes definitions. The inset identifies key elements of the robot design and illustrates how each wing is independently driven by a separate flight muscle. (**C**) Roll torque is generated by flapping one wing with larger stroke amplitude than the other, inducing differential thrust forces (*7*). (**D**) Pitch torque is generated by moving the mean stroke angle of both wings forward or backward to offset the thrust vector away from the center of mass. This mimics a method observed in *Drosophila* (*17*). (**E**) To generate yaw torques, the robot influences wing drag forces by cyclically modulating stroke velocity in a "split-cycle" scheme (*25*, *30*) rather than tilting the stroke plane (*10*) or altering wing angle of attack (*18*) as have been observed in *Drosophila*. A difference in stroke velocity between half-wing strokes results in an imbalanced drag force per stroke cycle—the higher velocity half-stroke (black arrow) produces greater drag force. By modulating magnitude and direction of this mean drag force on both wings, a net yaw torque is generated. The black and gray arrows correspond to arrows in (F). (**F**) The effect of stroke velocity on a wing's drag force. Black lines indicate the wings' position and pitch angle at temporally equidistant points within the stroke cycle. The red arrows indicate the instantaneous drag force on the wing.

remains at resonance (*20*). This is analogous to behavior observed in other flying insects (*1*, *24*).

The robotic fly departs from earlier single-actuator designs (*19*) by powering each wing with a separate, identical flight muscle to actuate two wings independently (*20*). This enables it to exert control torques about all three body axes (Fig. 1, B to F). The dual-actuator design weighs 80 mg, has a wingspan of 3 cm, and is capable of generating >1.3 mN of lift force (*20*). We measured power consumption of the robot to be 19 mW [which is consistent with similarly sized insects (*1*)], drawn from an offboard power source via a wire tether. If we implement onboard power with current technologies, we estimate no more than a few minutes of untethered, powered flight (*25*). Long duration power autonomy awaits advances in small, high-energy-density power sources.

Modulation of thrust force and three (approximately orthogonal) body torques (*25*) permits the robot to be controllable in unconstrained flight. To achieve stable flight, we must implement an active flight controller because, similar to flying insects, the dynamics of our insect-scale vehicle are fast and unstable (*4*, *26*), posing a difficult controller design problem reminiscent of the control of fighter jets (*27*) but analogously promising the potential for high-performance maneuverability. Sensing and controller computation are performed off-board, and power and control signals are sent to the robot via a wire tether consisting of four bundled, 51-gauge copper wires. The low mass (5 mg) and highly variable conformation of the tether suggests that it does not have a meaningful impact on the stability of the robot (*25*).

To sense the state of the robotic fly, we operated the robot in a virtual volume defined by an external array of motion-capture cameras; position and orientation were estimated by observing retroreflective tracking markers mounted on the robot (Fig. 2B). In contrast, flies appear to use dorsally oriented ocelli to estimate orientation and halteres to measure angular rates (*3*, *14*). Taking into account the sampling frequency of the motion-capture system (500 Hz), the latency of the computation, and the phase shift caused by the dynamics of the thorax, we estimated the total latency of the robotic fly's sensorimotor system to be ~12 ms (*25*). This proved to be sufficiently high bandwidth for the fast rotational dynamics of our insect-sized vehicle and is comparable with the 10-ms latency measured in the neuromotor reflexes of *Drosophila* (*15*).

Each robotic fly has distinct, unpredictable variations and asymmetries despite best efforts in manufacturing precision. Key system properties such as the flight muscle-thorax-wing system

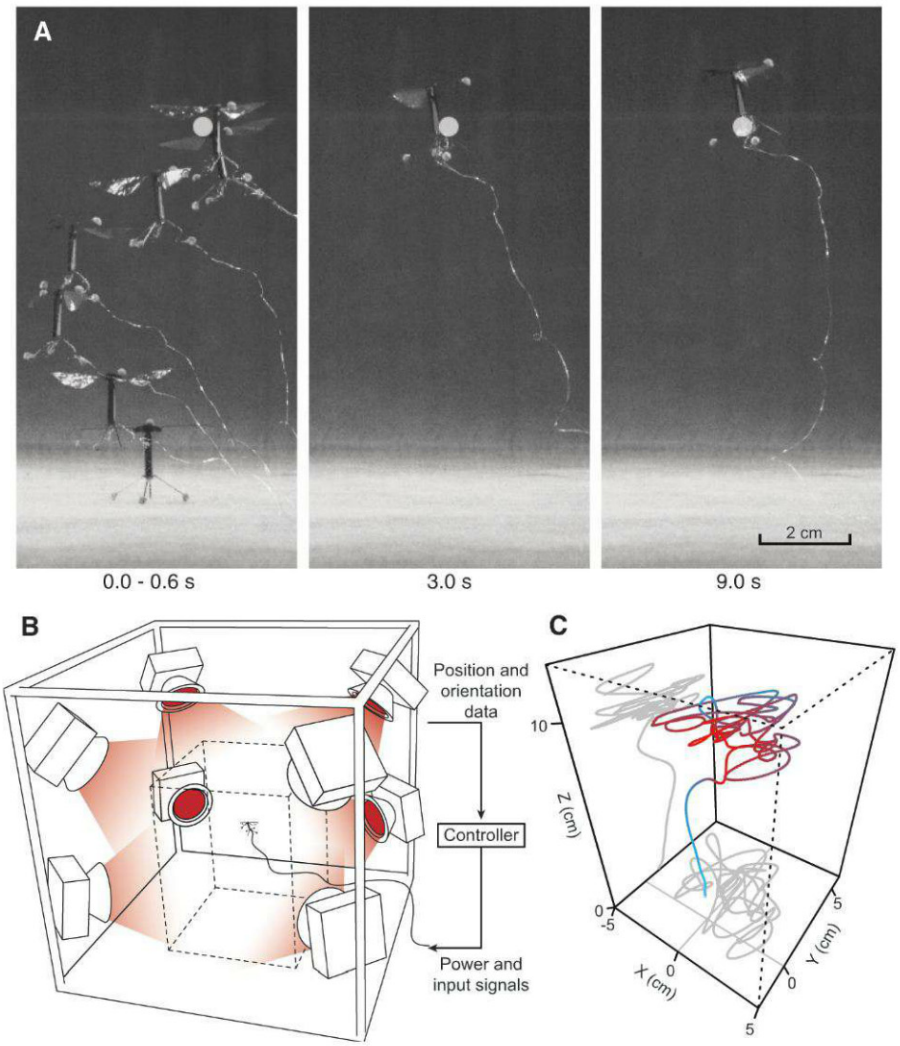


Fig. 2. Controlled takeoff and hovering of the robotic fly. (**A**) Frames from movie S1 at various times during the robot's flight. (Left) Strobed positions at 0.1-s intervals. The white dot indicates the desired hovering location. (**B**) Eight infrared, motion tracking cameras observe the positions of retroreflective markers attached to the robot in order to estimate its position and orientation in space with low latency. Position estimates are transmitted to the controller computer, which computes the control signals and sends them to the robot via a wire tether. (**C**) Three-dimensional reconstruction of the hovering flight trajectory in movie S1. Target position is 10 cm above ground. Line color gradient indicates distance from the target point, with red indicating closer proximity.

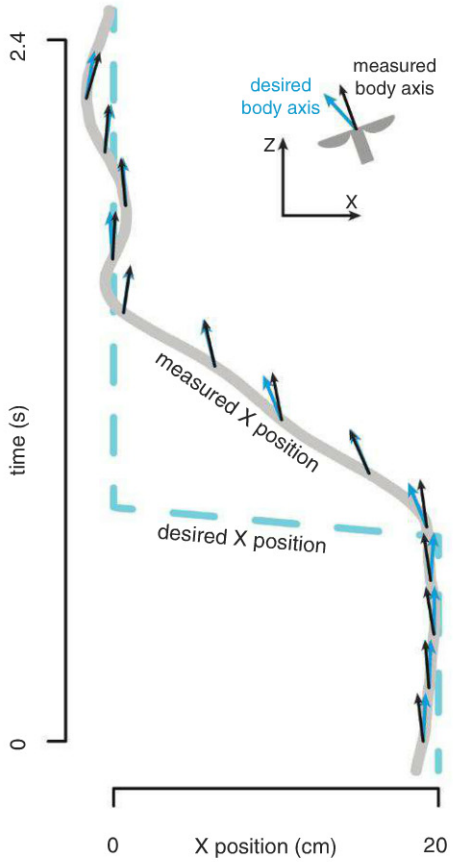


Fig. 3. The robotic fly executes a lateral maneuver (movie S2). During the maneuver, the altitude remains roughly constant. With time plotted on the vertical axis, the orientation of the robot's body axis, as projected onto the *x-z* plane, tilts to generate lateral forces in response to a change in the desired lateral position (dashed line). Robot orientation vectors are shown at temporally equidistant points 0.22 s apart so as to show the dynamics of the stabilization response and are not drawn to scale.

dynamics; exact magnitude, direction, and point of action of aerodynamic forces from wing trajectories; and the necessary compensation for torque biases from manufacturing inconsistencies are difficult to measure with commercially available sensors because of the small scale of the robot. Using a combination of custom-built sensors (*20*), theoretical models (*23*, *28*), and system identification during flight tests, we estimated these properties to aid in controller design (*25*).

The flight controller design consists of three distinct modules controlling body attitude, lateral position, and altitude and is subject to the constraints of the mechanical system; the stroke planes of the wings—and thus the direction of their time-averaged thrust vectors—essentially remain fixed with respect to the robot's body axis. To stay aloft, the robot must maintain a nominally upright orientation via stabilizing body torques so that its net thrust vector compensates for gravity. To induce lateral forces, the robot must reorient the body so that the net thrust vector takes on a lateral component (Fig. 3). This mimics the body tilt behavior observed in flies (*10*, *16*). Therefore, the body attitude controller module is critical. We use a Lyapunov function to derive an attitude control law that is asymptotically stable, given several simplifying assumptions (*25*). The control law consists of a proportional term that accounts for the error from a reference orientation and a derivative term that opposes angular velocity, providing rotational damping. The lateral position controller module operates by calculating the necessary reference orientation for the body attitude controller module to produce the appropriate lateral force component.

The altitude controller does not rely on information about body attitude; it is based on a linearization of the robot's dynamics at hover and assumes the system is always at an upright orientation. This decoupling of the controller allows for reduced constraints on the more sensitive attitude and lateral position controllers (*25*). In practice, the robot effectively maintains altitude because the body attitude does not deviate substantially from the nominal upright orientation, even when generating compensatory lateral forces.

In flight tests, the robotic fly demonstrated stable hovering about a fixed point with position errors on the order of one body length around the target position (Fig. 2, A and C, and movie S1), sustaining flights for longer than 20 s without ever approaching a crash. It also demonstrated lateral flight maneuvers, alternating between two fixed points in space by a switch of the target lateral position (Fig. 3 and movie S2).

Because of its scale and ability to perform stable, controlled flight, the robotic fly provides an alternative method for studying insect-scale, flapping-wing flight mechanics and flight control. For example, our flight data suggest that the robot's attitude-stabilization torques are highly dependent on information about angular velocities (Fig. 4). This coincides with the biological observation that haltere-mediated feedback is rate-dependent in *Drosophila* (*14*) and the similar finding in theoretical models of other flying insects (*26*). Additionally, flapping-wing flight experiences movement-based forces and torques that may be difficult to simulate in dynamically scaled fluid mechanics models. These dynamics, such as nonlinearities and cross-coupling of different degrees of freedom that arise during complicated flight maneuvers, could be measured with model fitting (*29*) or onboard sensors.

Even with a single actuated degree of freedom in each wing, the simplicity of the robot's mechanical design, in combination with scale-appropriate manufacturing and control strategies, is sufficient to enable controllable, insect-scale flapping-wing flight. The successful flight of the robotic fly demonstrates the feasibility of artificially approximating the flight apparatus of flying insects—particularly *Diptera*—in form and function and motivates future studies in miniaturized power, sensing, and computation technologies.

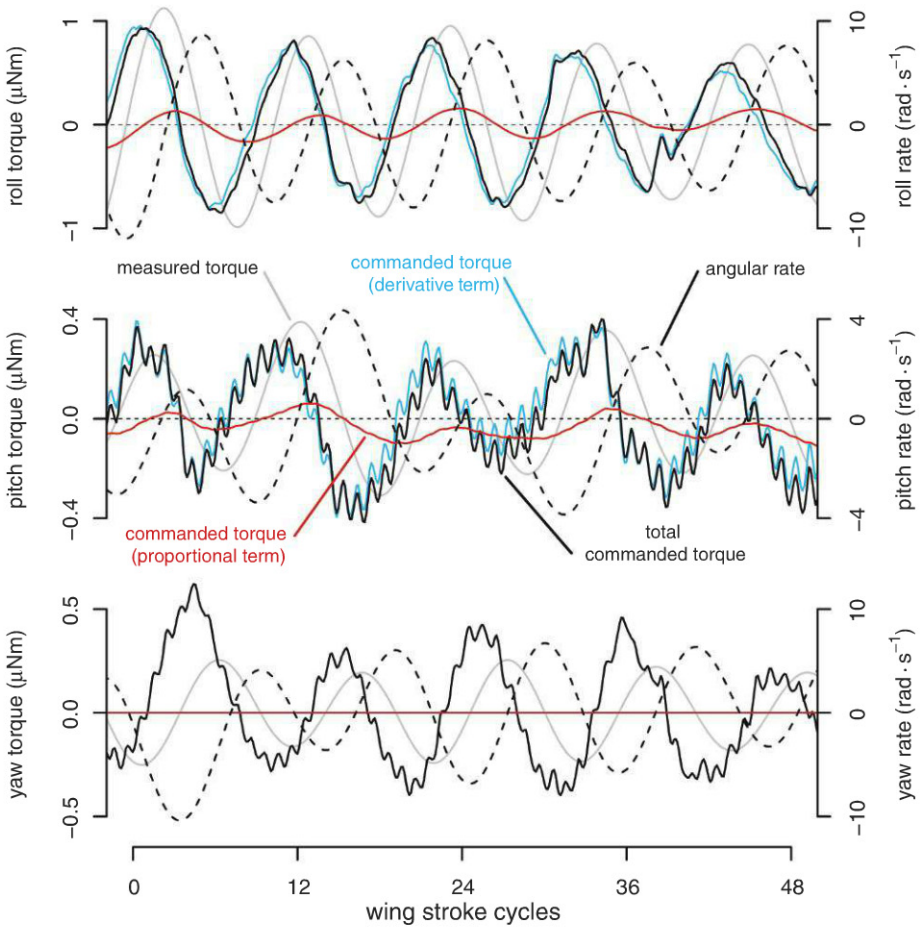


Fig. 4. Commanded torque magnitudes during an attitude-stabilized flight indicate that stabilization torques are dominated by the derivative term of the controller. In this study, lateral position control was turned off to emphasize torque commands from attitude stabilization. In the yaw torque plot, the proportional term does not contribute because, by design, only yaw rate is controlled (*25*). The commanded torque counters the measured angular rate with a delay of approximately two wing-stroke cycles. In all cases, the measured torque follows the commanded torque with a constant delay due to sensor latency and robot thoracic mechanics. Angular rate and measured torque (estimated from angular rate and the inertia of the robot) was low-pass filtered with a cutoff frequency of 60 Hz in order to reduce noise. The robot's flapping frequency is 120 Hz, and its effect can be observed as a higher-frequency oscillation superimposed on the commanded torque signals.

References and Notes

1. R. Dudley, *The Biomechanics of Insect Flight: Form, Function, Evolution* (Princeton Univ. Press, Princeton, NJ, 2002).
2. G. K. Taylor, *Biol. Rev. Camb. Philos. Soc.* **76**, 449 (2001).
3. G. Taylor, H. Krapp, *Adv. Insect Physiol.* **34**, 231 (2007).
4. C. Orlowski, A. Girard, *Prog. Aerosp. Sci.* **51**, 18 (2012).
5. M. H. Dickinson, F. O. Lehmann, S. P. Sane, *Science* **284**, 1954 (1999).
6. R. Wood *et al.*, *Int. J. Robot. Res.* **31**, 1292 (2012).
7. C. P. Ellington, *J. Exp. Biol.* **202**, 3439 (1999).
8. W. Trimmer, *Sens. Actuators* **19**, 267 (1989).
9. R. Wood, S. Avadhanula, R. Sahai, E. Steltz, R. Fearing, *J. Mech. Des.* **130**, 052304 (2008).
10. S. N. Fry, R. Sayaman, M. H. Dickinson, *Science* **300**, 495 (2003).
11. A. Ennos, *J. Exp. Biol.* **142**, 49 (1989).
12. A. Ennos, *J. Exp. Biol.* **140**, 161 (1988).
13. M. Dickinson, M. Tu, *Comp. Biochem. Physiol. A* **116**, 223 (1997).
14. M. H. Dickinson, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **354**, 903 (1999).
15. L. Ristroph *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 4820 (2010).

16. C. T. David, *Physiol. Entomol.* **3**, 191 (1978).
17. J. M. Zanker, *Physiol. Entomol.* **13**, 351 (1988).
18. A. J. Bergou, L. Ristroph, J. Guckenheimer, I. Cohen, Z. J. Wang, *Phys. Rev. Lett.* **104**, 148101 (2010).
19. R. Wood, *IEEE Trans. Robot.* **24**, 341 (2008).
20. K. Ma, S. Felton, R. Wood, "Design, fabrication, and modeling of the split actuator microrobotic bee," 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura, Algarve, Portugal, 7–12 October 2012, pp. 1133–1140.
21. M. Karpelson, G. Wei, R. Wood, "A review of actuation and power electronics options for flapping-wing robotic insects," 2008 IEEE International Conference on Robotics and Automation, Pasadena, CA, USA, 19–23 May 2008, pp. 779–786.
22. R. Wood, E. Steltz, R. Fearing, *Sens. Actuators A Phys.* **119**, 476 (2005).
23. J. Whitney, R. Wood, *J. Fluid Mech.* **660**, 197 (2010).
24. D. L. Altshuler, W. B. Dickson, J. T. Vance, S. P. Roberts, M. H. Dickinson, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 18213 (2005).
25. Materials and methods are available as supplementary materials on *Science* Online.
26. B. Cheng, X. Deng, T. L. Hedrick, *J. Exp. Biol.* **214**, 4092 (2011).
27. M. Abzug, E. Larrabee, *Airplane Stability and Control: A History of the Technologies that Made Aviation Possible* (Cambridge Univ. Press, Cambridge, UK, 2005).
28. B. Finio, N. Pérez-Arancibia, R. Wood, "System identification and linear time-invariant modeling of an insect-sized flapping-wing micro air vehicle," 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, CA, USA, 25–30 September 2011, pp. 1107–1114.
29. W. Hoburg, R. Tedrake, System identification of post stall aerodynamics for UAV perching. *Proc. AIAA Infotech Aerospace Conf.* (AIAA, Seattle, WA, 2009).
30. M. Oppenheimer, D. Doman, D. Sigthorsson, Dynamics and control of a biomimetic vehicle using biased wingbeat forcing functions: Part I—Aerodynamic model. *Proc. 48th AIAA Aerospace Sci. Meeting* (AIAA, Orlando, FL, 2010).

Acknowledgments: This material is based on work supported by the National Science Foundation under grant CCF-0926148, the Office of Naval Research under grant N00014-08-1-0919, and the Wyss Institute for Biologically Inspired Engineering. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Supplementary Materials
www.sciencemag.org/cgi/content/full/340/6132/603/DC1
Materials and Methods
Supplementary Text
Figs. S1 and S2
Table S1
References (*31*, *32*)
Movies S1 and S2

22 October 2012; accepted 12 February 2013
10.1126/science.1231806

3D Reconstruction of the Source and Scale of Buried Young Flood Channels on Mars

Gareth A. Morgan,[1]* Bruce A. Campbell,[1] Lynn M. Carter,[2] Jeffrey J. Plaut,[3] Roger J. Phillips[4]

Outflow channels on Mars are interpreted as the product of gigantic floods due to the catastrophic eruption of groundwater that may also have initiated episodes of climate change. Marte Vallis, the largest of the young martian outflow channels (<500 million years old), is embayed by lava flows that hinder detailed studies and comparisons with older channel systems. Understanding Marte Vallis is essential to our assessment of recent Mars hydrologic activity during a period otherwise considered to be cold and dry. Using data from the Shallow Radar sounder on the Mars Reconnaissance Orbiter, we present a three-dimensional (3D) reconstruction of buried channels on Mars and provide estimates of paleohydrologic parameters. Our work shows that Cerberus Fossae provided the waters that carved Marte Vallis, and it extended an additional 180 kilometers to the east before the emplacement of the younger lava flows. We identified two stages of channel incision and determined that channel depths were more than twice those of previous estimates.

The majority of outflow channels on Mars are attributed to megafloods caused by the catastrophic release of groundwater. The most prominent outflow channels, located around the Chryse Basin, are >1000 km long and are estimated to be Hesperian [~3.7 to 3.1 billion years ago (Ga)] in age (*1*–*3*). Marte Vallis in Elysium Planitia is the largest of the young (late Amazonian: ~0.5 Ga to the present) outflow channels on Mars. The channel system extends over ~1000 km in length and ~100 km in width, making Marte Vallis comparable in scale to the Chryse basin channel systems. Young lava flows have fully embayed the most elevated portions of Marte Vallis, and as a consequence the fundamental characteristics of the channels, including their source, depth, and morphology are less well understood than those of the Hesperian channels (*4*), despite being over 2.6 billion years younger (*5*).

Two possible sources have been proposed for Marte Vallis: water flowing from the Athabasca Valles outflow channel in the west (*4*, *6*, *7*), possibly forming bodies of water such as the putative frozen central Cerberus sea (*8*); and water flowing from a now-buried section of Cerberus Fossae (*5*, *6*, *9*). It is impossible to resolve which of the above hypotheses are correct from investigations of the surface geology alone. Using data from the Shallow Radar (SHARAD) sounder (*10*, *11*) on the Mars Reconnaissance Orbiter, we present a tomographic visualization of the buried Marte Vallis channels (*12*).

All 58 SHARAD tracks covering the uppermost reaches of Marte Vallis [as identifiable in Mars Orbiter Laser Altimeter (MOLA) gridded data] display multiple reflecting horizons (Fig. 1 and fig. S1). From mapping of the spatial distribution of SHARAD subsurface returns (Figs. 1 and 2), three distinct reflectors have been identified. Two of these reflectors are found extensively across the study area and occupy different depth ranges (*13*), referred to here as L1R (the shallower reflector) and L2R (the deeper reflector). The third reflector, R3, is located only in the southern portion of the region (Fig. 1C).

Radargrams reveal that the northern and southern termini of the R3 reflector dip upward and reconnect with the surface, delineating a discrete facies boundary (Fig. 1A and fig. S1). The R3 reflector is located exclusively below a mapped unit of young volcanics, ACy [(*5*), also mapped as AEc_3 by (*14*)] (Fig. 2C and fig. S2C). This unit is interpreted to be formed of voluminous lava flows <230 million years old (*5*, *14*), suggesting that R3 represents the base of a distinct surficial flow. The bases of young lava flows have also been identified by SHARAD west of Ascraeus Mons (*15*). The northern boundary of R3 shows strong spatial correlation with the boundary between ACy and the older unit ACo [>500 million years old (*5*)] (Fig. 2), implying that the lava embayed the preexisting ACo surface south of Cerberus Fossae and flowed toward the northeast (the dominant slope direction of the present surface). The northern portion of the R3 reflector exhibits prominent depressions, delineating subsurface channels (Fig. 1A). These channel features are ~20 km wide and extend for at least 50 km in a northeast direction. Seen in plan form, the channel features begin abruptly adjacent to one another along an orientation trending from northwest to southeast (Fig. 2C and fig. S3).

We interpret these features to be the highest elevated channels of Marte Vallis (Fig. 3), implying that the lava flow whose base is defined by R3 infilled the channels as the lavas flowed to the northeast. This indicates that the erosion of the outflow channel cut into the original underlying surface of unit ACo before the emplacement of the younger ACy lavas (fig. S4). This sequence of events confirms the young age of Marte Vallis and places the channel formation between the emplacement of units ACo and ACy [10 to 500 million years ago (Ma)], in agreement with (*5*).

The L1R and L2R reflectors are found extensively across the study region, suggesting that they represent regional boundaries between three

[1]Center for Earth and Planetary Studies, Smithsonian Institution, Washington, DC, USA. [2]NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA. [3]Jet Propulsion Laboratory, Pasadena, CA, USA. [4]Planetary Science Directorate, Southwest Research Institute, Boulder, CO, USA.

*Corresponding author. E-mail: morganga@si.edu

bedrock facies (Figs. 1 and 2). Further inspection of the radargrams reveals that both reflectors are punctuated by incisions of varying width (Fig. 1, B and C). These incisions are not random but spatially align to reveal complex networks. We interpret the spatial location of the incisions to represent where channels have been eroded through either one or both of the two bedrock boundaries delineated by L1R and L2R. A similar methodology has been applied to map out buried flood channels on Earth through the use of seismic profiles (*16*). No incisions are observable in L1R and L2R below unit ACo, and most SHARAD tracks reveal reflector incisions that spatially correlate with the boundary of ACo and ACy (Fig. 2). This indicates that the facies boundaries are systematically correlated with ACo. Our interpretation is further corroborated by the surface morphology exhibiting streamlined features (which represent bedrock "islands" between the channels) that are spatially correlated with isolated patches of the bedrock reflectors (Fig. 1, B and C, and Fig. 2). We argue that the teardrop-shaped hills and associated reflectors are remnant sections of the older ACo plains isolated by the erosional formation of the Marte Vallis channels (such streamlined features characterize most outflow channels on Mars) before infilling of the channels by the young ACy lavas.

The channels identified in R3 align with the truncation of the L1R and L2R reflectors (fig. S3), implying that they are part of the same channel network. Many martian outflow channels are sourced in chaos terrain, interpreted to be the result of subsidence from the rapid evacuation of groundwater (*17*). The abrupt opening of the channels as seen in R3 and the lack of any evidence for depressions at their source argue against the existence of chaos terrain at the head of Marte Vallis. Instead, the channel alignment at the source matches the orientation of the Cerberus Fossae graben system to the west (Fig. 3). This, and the absence of observable subsurface channels in the southern portions of R3, strongly support the hypothesis that Cerberus Fossae was the source of the Marte Vallis floods. This allows the position of this now-buried portion of Cerberus Fossae to be inferred and suggests that the fissure previously extended at least ~180 km to the east of its present surface expression (Fig. 3 and fig. S3).

The bed of the channel features cannot be mapped "downstream" beyond the areal extent of the R3 reflector. The sounder signal is increasingly attenuated as the thickness of lava above the buried channel bed increases, so we postulate that the R3 reflector may extend over a larger part of the study region at depths beyond the SHARAD signal-penetration limit. Quantifying the channel erosion north of Cerberus Fossae can be achieved by measuring the extent of the incisions into the L1R and L2R reflectors. The horizontal and vertical position of the reflectors on each side of an incision was recorded and combined with the elevations of the streamlined forms in the MOLA data. Connecting these erosional control points yields an approximation of the buried channel morphology for each radargram. Because of the density of the SHARAD coverage, it was possible to interpolate between neighboring tracks to produce a 3D model of the buried Marte Vallis channels (Fig. 4).

The SHARAD data reveal a complex channel system consisting of a broad ~40-km-wide main channel that is adjacent to a raised bench, 120 km in width and incised by anastomosing channels formed around four streamlined islands. Such morphology is consistent with the majority of martian outflow channels, and the scale is comparable to that of the main channel and associated perched tributaries of the prominent Ares Vallis outflow channel system (Hesperian in age) that flows into the Chryse Basin (*18*). The geomorphic configuration of Marte Vallis implies the the system experienced two phases of erosion: (i) to erode the islands and (ii) to cut the main channel below the level of the perched channels. This also implies that the initial stages of channel formation consisted of small-scale anastomosing patterns before flow was concentrated in a deeper, wider channel, leaving the central islands and smaller channels as perched remnants (Fig. 4, fig. S4, and supplementary text). Recent work citing the turbulent nature of the young Elysium Planitia lavas suggests that lava flows may have partially or fully eroded the outflow channels (*19, 20*). Although they are filled with later lavas, our tomographic models show that the Marte Vallis channels are morphologically similar to the circum-Chryse outflow channels, and thus could reasonably have been carved entirely by water.

Estimates of the depth of the channels from time-delay offsets to the surface echo are based on an assumed range of values for the permittivity of the material overlying the L1R and L2R

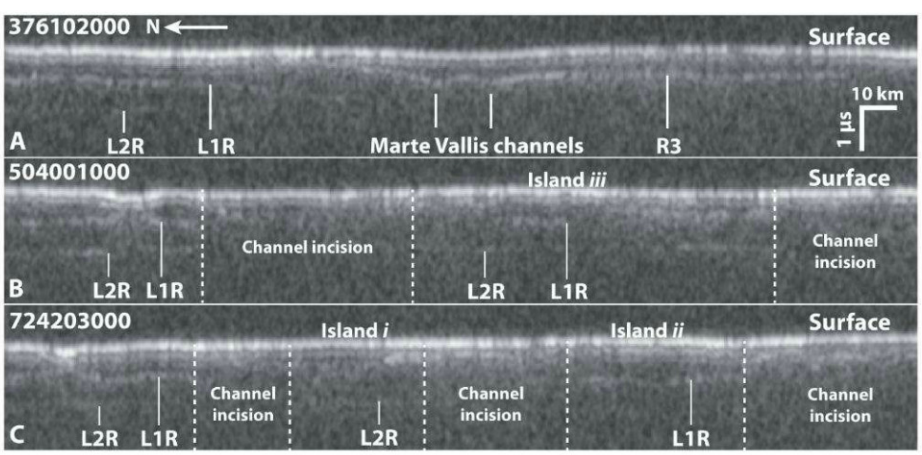


Fig. 1. SHARAD radargrams display the time delay of the echo signal against along-track distance. Subsurface contrasts in permittivity reflect the transmitted signal, producing the observed reflectors. Spatial coordinates of the radargram sections presented are as follows: (**A**) 376102000 (8.42°N, 175.61°E to 4.07°N, 175.08°E), (**B**) 504001000 (10.13°N, 177.81°E to 5.78°N, 177.28°E), (**C**) 724203000 (9.45°N, 176.79°E to 5.1°N, 176.26°E). The locations of the SHARAD tracks are presented in Fig. 2.

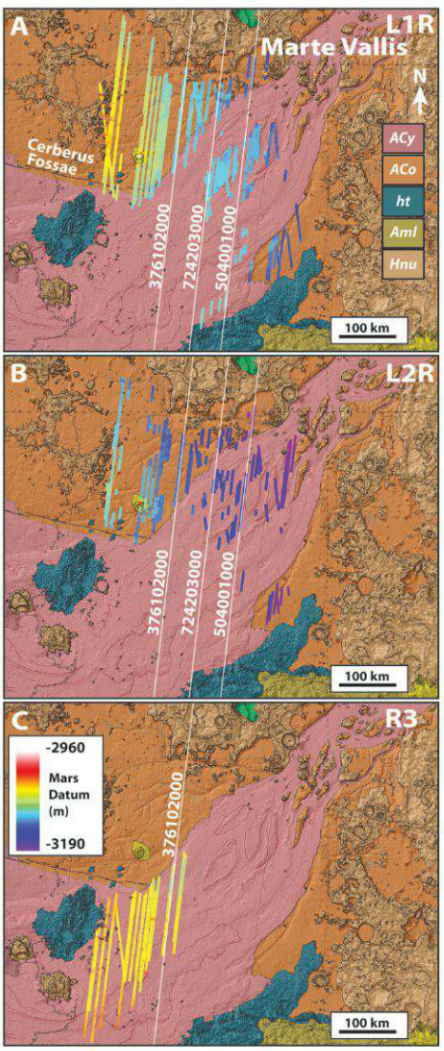


Fig. 2. Spatial distribution of L1R (A), L2R (B), and R3 (C) reflectors corrected for depth (*13*). The background shows the (*5*) geologic map above a hillshade image of MOLA gridded data (200 × vertical exaggeration). ht, hackly terrains; Aml, Medusae Fossae Formation; Hnu, undivided material forming hills and knobs. [Geologic map adapted from (*5*) with permission from Elsevier]

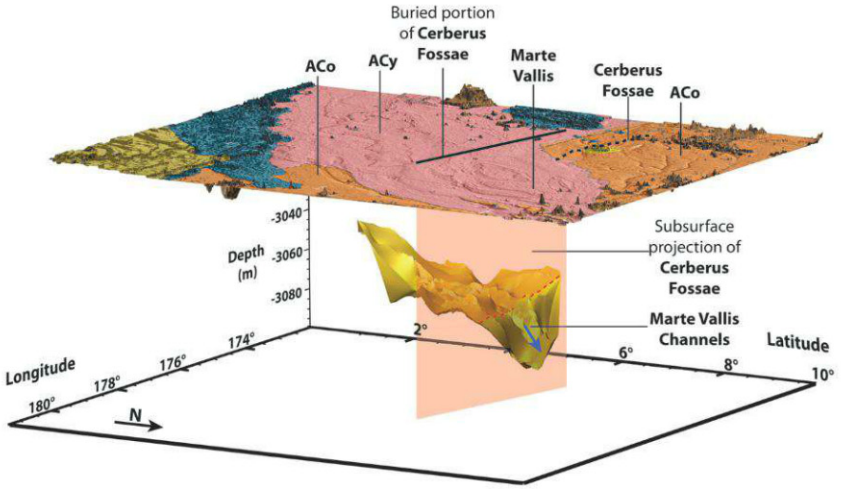


Fig. 3. 3D visualization of the R3 reflector [corrected for depth (*13*)] below the MOLA surface overlain with geologic units (*5*). The surface has been elevated and scaled by a factor of 1/150 for clarity. The Marte Vallis channels are visible as prominent depressions in the R3 reflector and begin abruptly along an orientation that aligns with the subsurface vertical projection of Cerberus Fossae. The blue arrow highlights the direction of flow in the channels. [Geologic map adapted from (*5*) with permission from Elsevier]

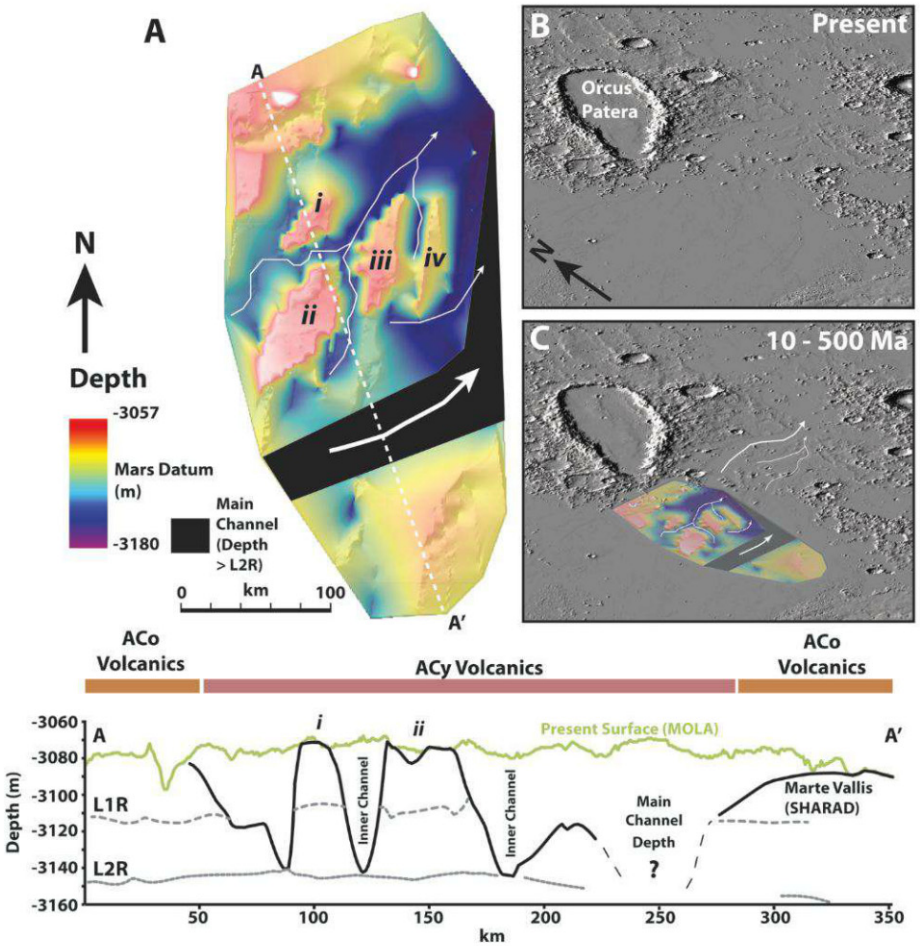


Fig. 4. (A) Tomographic model of the buried Marte Vallis channels and associated NW-SE topographic and tomographic profile (*13*). The model shows an ~40-km-wide main channel and adjacent teardrop-shaped islands. (**B** and **C**) show a perspective view of eastern Elysium Planitia at present (B) and before embayment by young lavas (C). White arrows represent the direction of flow.

reflectors (*21*). The permittivity of these materials is uncertain, but the typical range for dry geological materials is 3 to 8 (*22*, *23*), with lower values leading to greater depths below the surface for any given time delay. For example, the depth of the perched channel between islands ii and iii is constrained between the incised L1R and the continuous L2R reflectors. Applying a permittivity range of 8 to 3 yields a depth range of 35 to 56 m (L1R), to 67 to 110 m (L2R), respectively. With regard to the main channel, the depth estimate (between island i and the south bank) is 69 to 113 m. This is comparable to the depth of incision of the largest known megaflood on Earth, the Missoula floods, responsible for carving the Channeled Scabland of the northwestern United States (*24*). However, the Marte Vallis estimate represents a minimum value, because the L2R reflector has been fully bisected by this channel. The depths of the channels are at least double the previous maximum estimates for Marte Vallis of 40 m (*5*), demonstrating that the scale of the floods has been underestimated.

References and Notes

1. D. H. Scott, K. L. Tanaka, *Geological Map of the Western Equatorial Region of Mars. U.S. Geological Survey Miscellaneous Map I-1802-A* (U.S. Geological Survey, Washington, DC, 1986).
2. K. L. Tanaka, *J. Geophys. Res.* **91**, E139 (1986).
3. S. Rotto, K. L. Tanaka, *Geologic/Geomorphic Map of the Chryse Planitia Region of Mars. U.S. Geological Survey Miscellaneous Investigations Map I-2441* (U.S. Geological Survey, Washington, DC, 1995).
4. D. C. Berman, W. K. Hartmann, *Icarus* **159**, 1 (2002).
5. J. Vaucher *et al.*, *Icarus* **204**, 418 (2009).
6. D. M. Burr, J. A. Grier, A. S. McEwan, L. P. Keszthelyi, *Icarus* **159**, 53 (2002).
7. E. R. Fuller, J. W. Head, *J. Geophys. Res.* **107**, 5081 (2002).
8. J. B. Murray *et al.*; HRSC Co-Investigator Team, *Nature* **434**, 352 (2005).
9. J. B. Plescia, *Icarus* **164**, 79 (2003).
10. R. Seu *et al.*, *Planet. Space Sci.* **52**, 157 (2004).
11. R. Seu *et al.*, *J. Geophys. Res.* **112**, E05S05 (2007).
12. Materials and methods are available as supplementary materials on *Science* Online.
13. The time delay offset between a reflector and the surface echo within a radargram can be converted to depth. This depth estimation is inversely proportional to the square root of the assumed permittivity (ε′) of the material above the reflector. Figures 2 to 4 assume an ε′ of 8, consistent with dense, dry, terrestrial lava flows. Applying a lower value for ε′ would provide a greater depth estimate for each of the reflectors.
14. K. L. Tanaka, J. A. Skinner, T. M. Hare, *Geologic Map of the Northern Plains of Mars, U.S. Geological Survey Scicientific Investigations Map 2888* (U.S. Geological Survey, Washington, DC, 2005).
15. L. M. Carter *et al.*, *Geophys. Res. Lett.* **36**, L23204 (2009).
16. C. S. Fulthorpe, J. A. Austin Jr., *Geology* **32**, 1013 (2004).
17. M. Carr, *J. Geophys. Res.* **84**, 2995 (1979).
18. N. H. Warner, S. Gupta, J.-P. Muller, J.-R. Kim, S.-Y. Lin, *Earth Planet. Sci. Lett.* **288**, 58 (2009).
19. W. L. Jaeger *et al.*, *Icarus* **205**, 230 (2010).
20. D. W. Leverington, *Geomorphology* **132**, 51 (2011).
21. B. A. Campbell *et al.*, *J. Geophys. Res.* **113**, E12010 (2008).
22. F. T. Ulaby *et al.*, *Microwave Dielectric Spectrum of Rocks* (Report 23817-1-T, Univ. of Michigan Radiation Laboratory, Ann Arbor, MI, 1988).
23. W. D. Carrier, G. R. Ohloeft, W. Mendell, Physical properties of the lunar surface, in *Lunar Sourcebook*, G. H. Heiken, D. T. Vaniman, B. M. French, Eds. (Cambridge Univ. Press, New York, 1991), pp. 457–567.

24. V. R. Baker, D. Nummedal, Eds., *The Channeled Scabland; a Guide to the Geomorphology of the Columbia Basin* (NASA, Washington, DC, 1978).

Acknowledgments: C. Chazen and T. Watters provided very helpful reviews of earlier versions of the paper. Comments and suggestions by four anonymous referees were extremely beneficial. Thanks are also extended to P. Russell for his technical assistance. Funding for this work was provided by the NASA Mars Reconnaissance Orbiter Project. The SHARAD data and MOLA altimetry are available through NASA's Planetary Data system.

Supplementary Materials
www.sciencemag.org/cgi/content/full/science.1234787/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S4
Table S1
Reference (*25*)

4 January 2013; accepted 18 February 2013
Published online 7 March 2013;
10.1126/science.1234787

Structural Basis for Molecular Recognition at Serotonin Receptors

Chong Wang,[1]* Yi Jiang,[1,2]* Jinming Ma,[1,3]* Huixian Wu,[1] Daniel Wacker,[1] Vsevolod Katritch,[1] Gye Won Han,[1] Wei Liu,[1] Xi-Ping Huang,[4] Eyal Vardy,[4] John D. McCorvy,[4] Xiang Gao,[3] X. Edward Zhou,[3] Karsten Melcher,[3] Chenghai Zhang,[2,3] Fang Bai,[5] Huaiyu Yang,[6] Linlin Yang,[6] Hualiang Jiang,[6] Bryan L. Roth,[4] Vadim Cherezov,[1] Raymond C. Stevens,[1]† H. Eric Xu[2,3]†

Serotonin or 5-hydroxytryptamine (5-HT) regulates a wide spectrum of human physiology through the 5-HT receptor family. We report the crystal structures of the human 5-HT$_{1B}$ G protein–coupled receptor bound to the agonist antimigraine medications ergotamine and dihydroergotamine. The structures reveal similar binding modes for these ligands, which occupy the orthosteric pocket and an extended binding pocket close to the extracellular loops. The orthosteric pocket is formed by residues conserved in the 5-HT receptor family, clarifying the family-wide agonist activity of 5-HT. Compared with the structure of the 5-HT$_{2B}$ receptor, the 5-HT$_{1B}$ receptor displays a 3 angstrom outward shift at the extracellular end of helix V, resulting in a more open extended pocket that explains subtype selectivity. Together with docking and mutagenesis studies, these structures provide a comprehensive structural basis for understanding receptor-ligand interactions and designing subtype-selective serotonergic drugs.

The neuromodulator serotonin [5-hydroxytryptamine (5-HT)] is essential for diverse functions at nearly every organ system in the human body (*1*–*4*). The activity of 5-HT is mediated through activation of members of a large family of 5-HT receptor proteins that can be grouped into seven subfamilies (5-HT$_{1\text{-}7}$) on the basis of sequence homology and signaling mechanisms (*5*). Except for the 5-HT$_3$ receptor, which is a ligand-gated ion channel, the other 12 members are heterotrimeric guanine nucleotide binding protein (G protein)–coupled receptors (GPCRs). The serotonergic system is a target of many widely prescribed drugs, including atypical antipsychotics, antimigraine medications, anxiolytics, and antidepressants (*1*), and the recently approved antiobesity medication lorcaserin (*6*, *7*). However, clinical use of several serotonergic drugs caused unexpected side effects arising from off-target interactions with 5-HT receptor subtypes and related receptors for biogenic amine (*1*, *4*, *8*, *9*).

The 5-HT$_{1B}$ receptor couples to G protein alpha subunits G_i or G_o and is widely expressed in the brain and the cardiovascular system. In the CNS, the 5-HT$_{1B}$ receptor functions as an inhibitory presynaptic receptor to modulate the release of 5-HT and many other neurotransmitters (*1*, *2*). The 5-HT$_{1B}$ receptor is a primary molecular target for the antimigraine drugs ergotamine (ERG) and dihydroergotamine (DHE), which are efficacious 5-HT$_{1B}$ receptor agonists (*10*). Off-target activation of the related 5-HT$_{2B}$ receptor is responsible for the valvulopathic activity of many approved drugs and is the main reason for their withdrawal (*9*–*12*). We report two crystal structures of the human 5-HT$_{1B}$ receptor bound to the full agonists, ERG and DHE (tables S1 and S2). Comparison with the structure of the human 5-HT$_{2B}$ receptor bound to ERG (*13*) reveals critical structural determinants for ligand recognition and subtype selectivity and provides a structural rationale for designing safer and more effective serotonergic drugs.

Crystallization studies of the 5-HT$_{1B}$ receptor were done with engineered constructs, 5-HT$_{1B}$-1 and 5-HT$_{1B}$-2 (*14*), which crystallized with ERG and DHE at resolutions of 2.7 Å and 2.8 Å, respectively. Due to the high similarity between these two structures (fig. S2), for brevity we focus on the structure of the 5-HT$_{1B}$-1/ERG complex for analysis and discussion of key structural features for ligand recognition and selectivity in 5-HT$_{1B}$ versus 5-HT$_{2B}$ receptors.

The main fold of the 5-HT$_{1B}$ receptor consists of a canonical seven-transmembrane (7TM) α-helical bundle (Fig. 1A). The extracellular loop 2 (ECL2) that partially covers the ligand binding pocket is stabilized by a C122$^{3.25}$-C199^{ECL2} disulfide bond, highly conserved in GPCRs. Part of the N terminus folds on top of the binding pocket, where Y40 forms hydrogen-bond interactions with ligand binding residue D352$^{7.36}$ (fig. S5) (*15*, *16*). This feature suggests that the N terminus could have a role in ligand recognition in the 5-HT$_{1B}$ receptor by interacting with residues within the binding pocket.

The 5-HT$_{1B}$/ERG complex structure revealed a large ligand binding cavity defined by residues from helices III, V, VI, VII, and ECL2, comprising an orthosteric pocket embedded deep in the 7TM core and an extended binding pocket close to the extracellular entrance (Fig. 1). ERG adopts a binding mode with the ergoline ring system occupying the orthosteric binding pocket and the cyclic tripeptide moiety bound to the upper extended binding pocket (Fig. 2C). In the orthosteric pocket, the ergoline scaffold is anchored through the salt-bridge interaction between its positively charged nitrogen and the carboxylate of D129$^{3.32}$, which is fully conserved in 5-HT and other monoamine receptors. The side chain of D129$^{3.32}$ is further stabilized by a hydrogen bond to the hydroxyl of Y359$^{7.43}$. Side chains of C133$^{3.36}$, I130$^{3.33}$, W327$^{6.48}$, F330$^{6.51}$, and F331$^{6.52}$ form a narrow hydrophobic cleft, which packs tightly against the nearly planar ergoline ring system. In addition, the indole N-H hydrogen forms a hydrogen bond with T134$^{3.37}$ (Fig. 2A). Comparison with the ERG-bound 5-HT$_{2B}$ receptor structure revealed that the orthosteric binding pockets in the two receptors are very similar, with the key interactions conserved (Fig. 2, A, D, and E). The only difference is observed in the region where residues from helix V contact ERG. Due to the lack of a side chain at G221$^{5.42}$, the side chain of F217$^{5.38}$ in the 5-HT$_{2B}$ receptor reaches into the ligand binding pocket and packs on top of the ERG indole ring; by comparison, the corresponding interaction in the 5-HT$_{1B}$ receptor occurs between the side chain of S212$^{5.42}$ and ERG, whereas Y208$^{5.38}$ does not interact with the ligand due to the outward shift of helix V (Fig. 2, A and D). Substantial differences are observed in the extended

[1]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA. [2]Van Andel Research Institute/Shanghai Institute of Materia Medica Center, CAS-Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China. [3]Laboratory of Structural Sciences, Van Andel Research Institute, 333 Bostwick Avenue Northeast, Grand Rapids, MI 49503, USA. [4]National Institute of Mental Health Psychoactive Drug Screening Program, Department of Pharmacology and Division of Chemical Biology and Medicinal Chemistry, University of North Carolina Chapel Hill Medical School, Chapel Hill, NC 27599, USA. [5]Department of Engineering Mechanics, State Key Laboratory of Technology, Faculty of Chemical, Environmental, and Biological Science and Technology, Dalian University of Technology, Dalian 116023, China. [6]State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China.

*These authors contributed equally to this work.
†Corresponding author. E-mail: eric.xu@vai.org (H.E.X); stevens@scripps.edu (R.C.S.)

binding pockets. The extended binding pocket of the 5-HT_{1B} receptor is broader than that of the 5-HT_{2B} receptor (Fig. 1, B, C, and D) due to the 3.0 Å outward shift of the top of helix V. Moreover, in contrast to the relatively bulky $M218^{5.39}$ of the 5-HT_{2B} receptor, the corresponding residue of the 5-HT_{1B} receptor is a smaller threonine, which results in a further expansion of this pocket (Fig. 2B). Despite these differences in contact residues between the two subtypes, the cyclic tripeptide moiety of ERG maintains a similar overall orientation relative to the ergoline moiety, likely stabilized by an intramolecular hydrogen bond (Fig. 2, D and E). The conformations of the phenyl group of the ERG cyclic tripeptide do differ. It contacts $L347^{6.58}$ and $V348^{6.59}$ in the 5-HT_{2B} receptor, whereas in the 5-HT_{1B} receptor, it rotates to occupy a cavity close to $T209^{5.39}$ at helix V. Correspondingly, $M337^{6.58}$ turns into the pocket to contact with the phenyl ring of the ligand in the 5-HT_{1B} receptor. As shown in the accompanying paper, these differences in ERG interactions correlate with different functional states. At the 5-HT_{1B} receptor, ERG causes full activation, whereas ERG induces intermediate G protein activation and β-arrestin biased signaling at the 5-HT_{2B} receptor (*13*).

Mutations of several residues in the orthosteric binding site of the 5-HT_{1B} receptor, including the conserved $D129^{3.32}$, abolished the binding of the radioligand lysergic acid diethylamide (LSD) (table S4). The prototypical hallucinogen LSD has the same ergoline moiety as ERG, and these interactions within the orthosteric pocket appear to be the main driving force for binding of ergolines. In contrast, none of the tested mutations of the extended binding pocket residues considerably reduced the binding affinity of ERG, suggesting that the cyclic tripeptide moiety of ERG is accommodated without contributing substantially to the binding affinity. This extended binding site partially overlaps with the sites inferred in interaction with allosteric modulators in the M_2 muscarinic acetylcholine receptor (*17*, *18*), suggesting its potential role in mediating allosteric modulations at 5-HT receptors (*13*).

Alignment of all human 5-HT GPCR sequences shows that the residues in the orthosteric binding pockets are much more conserved than those in the extended binding pockets (fig. S7). This likely reflects an evolutionary pressure to maintain the structure of the orthosteric binding pocket for recognition of the endogenous ligand 5-HT (*19*). Molecular docking of 5-HT into the orthosteric binding pocket of the 5-HT_{1B} and 5-HT_{2B} receptors revealed important residues involved in the recognition of 5-HT (Fig. 3A), many of which have been previously implicated by site-directed mutagenesis and molecular modeling studies (*20*, *21*). 5-HT shares a common chemical substructures with ergolines (fig. S4) and recognizes the 5-HT receptors in a similar manner as ERG and DHE. $D^{3.32}$ forms a salt bridge with the positively charged amino group of 5-HT, whereas $T^{3.37}$, which is highly conserved in 5-HT receptors and most other aminergic GPCRs, forms a hydrogen bond with the indole N-H hydrogen. This hydrogen bond appears to be important for the recognition of 5-HT by 5-HT_{1B} and 5-HT_{2B} receptors because mutating $T^{3.37}$ to alanine reduces the affinity of 5-HT at both receptors by a factor of more than

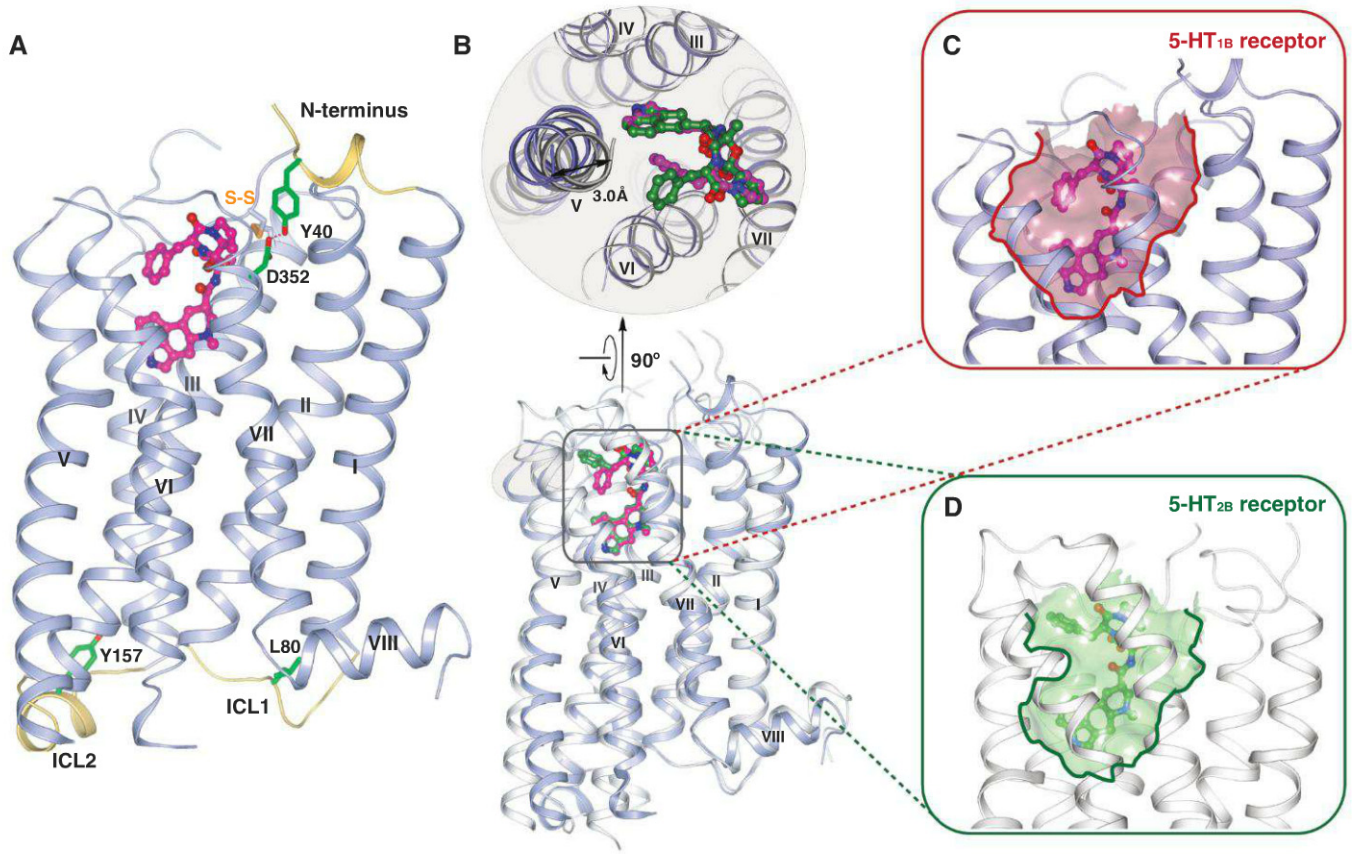


Fig. 1. Overall architecture of the 5-HT_{1B} receptor bound to ERG and comparison of the ligand binding pocket shapes of the 5-HT_{1B} receptor and the 5-HT_{2B} receptor. (**A**) The 5-HT_{1B} receptor is shown as a light blue colored ribbon cartoon; the N terminus, ICL1, and ICL2 are highlighted in yellow. Ligand ERG is colored magenta. The disulfide bond between C122 and C199 is shown as orange sticks. The Y40 and $D352^{7.36}$ side chains, which mediate the interaction between the N terminus and the ligand binding pocket, are shown as green sticks, with a dashed line indicating the hydrogen bond interaction. $L80^{ICL1}$ and $Y157^{ICL2}$ interact with residues of the 7TM bundle stabilizing the local structures of ICL1 and ICL2, respectively. (**B**) Shown at the bottom is the superposition of the 5-HT_{1B}/ERG structure (light blue) and the 5-HT_{2B}/ERG structure (white). The ligands are colored magenta for the 5-HT_{1B} receptor and green for the 5-HT_{2B} receptor. The top panel shows an extracellular view of the ligand binding sites. The arrow indicates a 3.0 Å shift (distance measured between the α carbons of $T209^{5.39}$ in the 5-HT_{1B} receptor and $M218^{5.39}$ in the 5-HT_{2B} receptor) at the extracellular end of helix V. (**C** and **D**) The surface representation of the ligand binding pockets of the 5-HT_{1B} receptor and the 5-HT_{2B} receptor are shown in transparent pink and transparent green, respectively.

10 (table S4). The indole ring points toward residues on helix V: S212$^{5.42}$ and A216$^{5.46}$ at the 5-HT$_{1B}$ receptor, G221$^{5.42}$ and A225$^{5.46}$ at the 5-HT$_{2B}$ receptor. This ligand-receptor interface is less polar in the 5-HT receptor family compared with those of other biogenic amine receptors (table S5), thereby perfectly matching the property of the 5-HT indole head group, which is less polar than those of other aminergic receptor native agonists, such as epinephrine, dopamine, and histamine.

LSD promiscuously binds to 5-HT receptors: It has agonist activity at most 5-HT receptors (table S6), whereas it is a potent antagonist at 5-HT$_{7A}$ receptor (*13*). Docking of LSD into the 5-HT$_{1B}$ and 5-HT$_{2B}$ receptor structures suggests that binding of the LSD ergoline moiety in the orthosteric binding pocket is essentially identical to that of the ERG ergoline moiety (Fig. 3B). Many single point mutations within the orthosteric binding pocket reduce or abolish the binding of LSD at both 5-HT$_{1B}$ and 5-HT$_{2B}$ receptors (table S4). The conservation of the orthosteric binding pocket provides an atomic-level explanation for the observed promiscuous interactions between 5-HT receptors and drugs like LSD.

Triptans are 5-HT analogs that are among the most frequently prescribed antimigraine medications that act primarily through 5-HT$_{1B}$ and 5-HT$_{1D}$ receptors (*1*, *2*, *22*). A common structural feature of triptans is the large substitution group at the 5′ position of the indole ring. Functional assays indicate that triptans act as potent agonists of the 5-HT$_{1B}$ receptor but not the 5-HT$_{2B}$ receptor (table S7). To address the structural determinants of this selectivity, we performed docking simulations at both 5-HT$_{1B}$ and 5-HT$_{2B}$ receptors (fig. S8). Whereas triptans were well accommodated in the 5-HT$_{1B}$ receptor binding pocket, the narrower extended binding pocket in the helix V region of the 5-HT$_{2B}$ receptor forced the ligands to adopt unfavorable positions (Fig. 3C and fig. S8C). Among the tested triptans, donitriptan and eletriptan showed relatively higher potency in G$_q$-mediated signaling at the 5-HT$_{2B}$ receptor (table S7). These two triptans have longer and more flexible linkers between 5′ substituents and the indole ring, en-

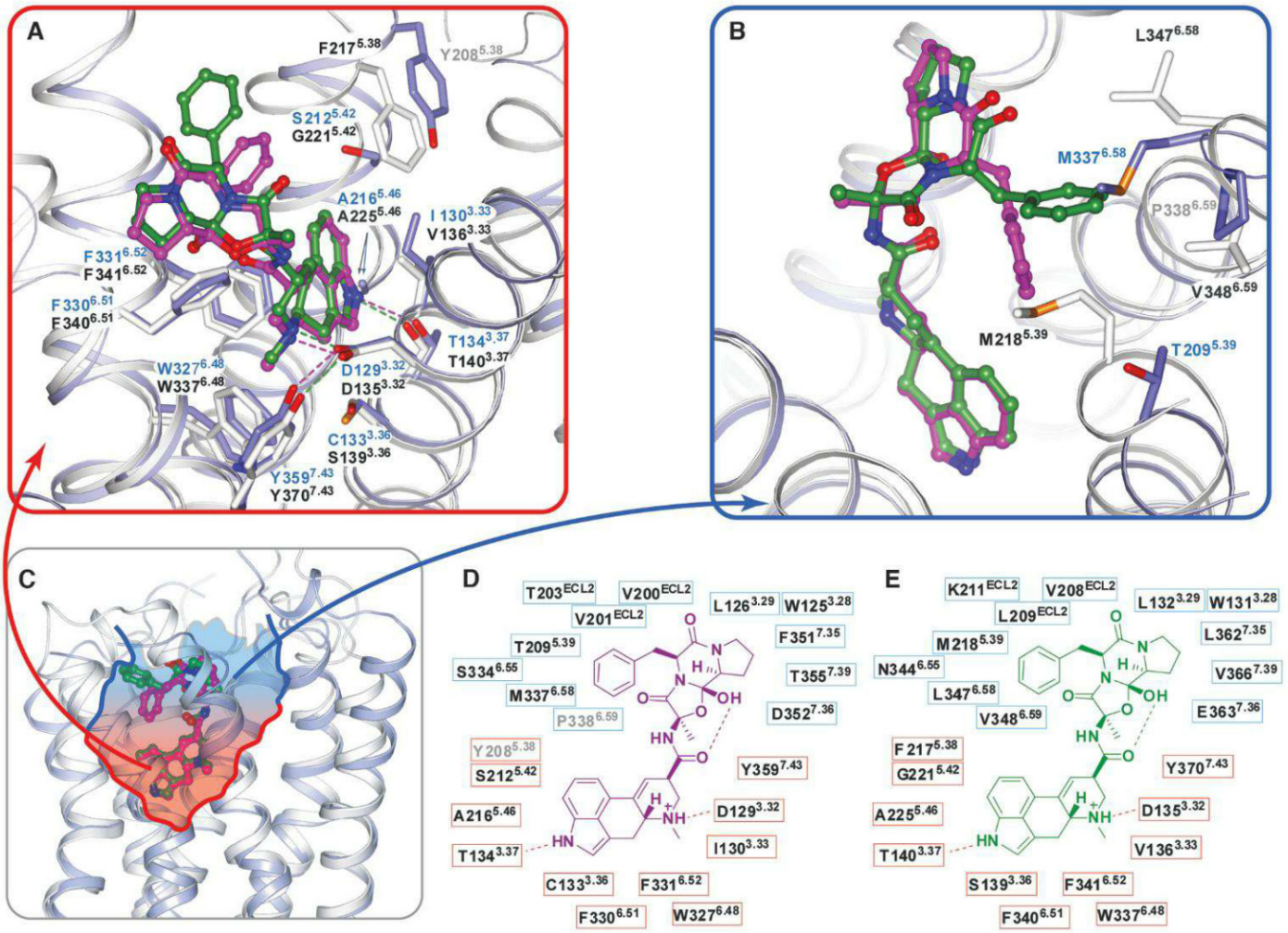


Fig. 2. Comparison of ligand-receptor interactions in 5-HT$_{1B}$/ERG and 5-HT$_{2B}$/ERG structures. (**A** to **C**) Superposition of the ligand binding pockets of 5-HT$_{1B}$ receptor (light blue) and 5-HT$_{2B}$ receptor (white). The carbons of ligand ERG in the 5-HT$_{1B}$ and the 5-HT$_{2B}$ receptor structures are shown as magenta and green, respectively. (A) Residues forming the orthosteric binding sites are shown as sticks and labeled in blue for 5-HT$_{1B}$ receptor and black for 5-HT$_{2B}$ receptor. The salt-bridge interactions between D$^{3.32}$ and ERG—as well as the hydrogen bond interactions between T$^{3.37}$ and ERG, and between Y$^{7.43}$ and D$^{3.32}$—are shown as red dashed lines for 5-HT$_{1B}$ receptor and green dashed lines for 5-HT$_{2B}$ receptor. (B) Substantial differences in the extended ligand binding pockets result in different conformations of the ligand ERG. (C) The overall binding pockets with the orthosteric and the extended ligand binding sites shaded red and blue, respectively. (**D** and **E**) Diagram representation of ligand interactions in the binding pockets of 5-HT$_{1B}$ and 5-HT$_{2B}$ receptors, respectively. Intramolecular hydrogen bonds within ERG are indicated as dashed lines. Residues in the orthosteric binding pockets are shown in red boxes, and extended binding pocket residues are shown in blue boxes. The hydrogen bond interaction between T$^{3.37}$ and ERG and the salt-bridge interactions between D$^{3.32}$ and ERG are indicated by red dashed lines. In (A), (B), and (D), Y208$^{5.38}$ and P338$^{6.59}$ of the 5-HT$_{1B}$ receptor, which do not interact with ERG, are labeled in gray.

abling them to fit better in the narrow binding pocket of the 5-HT_{2B} receptor. The $M218^{5.39}A$ mutation of the 5-HT_{2B} receptor, which increases the space in the binding pocket, substantially enhanced the potency of donitriptan and eletriptan (fig. S9), although their potency was still lower than that at the 5-HT_{1B} receptor. Thus, the broader opening near the extracellular end of helix V in the 5-HT_{1B} receptor appears to be important for selectively accommodating the 5′ substituents of triptan ligands, whereas the narrower pocket in the 5-HT_{2B} receptor shows reduced binding to these compounds.

Norfenfluramine is the active metabolite of fenfluramine, which is one of the two components of the infamous "Fen-Phen" antiobesity cocktail. The antiobesity effect of norfenfluramine occurs mainly through activation of the 5-HT_{2C} receptor (*1*), but due to its high potency and efficacy as a 5-HT_{2B} receptor agonist, it can cause life-threatening side effects, including pulmonary hypertension and heart valve disease (*9*, *23*–*25*). To elucidate the structural basis for norfenfluramine's subtype selectivity, we simulated its binding to the 5-HT_{1B} and 5-HT_{2B} receptor structures (Fig. 3D). Norfenfluramine tightly fit the orthosteric binding pocket of the 5-HT_{2B} receptor, with $F217^{5.38}$ and $M218^{5.39}$ forming a hydrophobic cap that interacted with the trifluoromethyl group. The $F217^{5.38}A$ mutation (table S8) and the $M218^{5.39}V$ mutation (*26*) both reduce the potency of norfenfluramine compared with that for the wild-type 5-HT_{2B} receptor. In the docking model of the 5-HT_{1B} receptor, these close contacts are missing; thus, the potency of norfenfluramine is reduced compared with that at the 5-HT_{2B} receptor (table S8).

Species-specific differences in the ligand binding properties of 5-HT_{1B} and other 5-HT receptors from rodents and humans have impeded the extrapolation of findings from animal models to humans (*27*, *28*). The rodent 5-HT_{1B} receptors, for instance, have a much higher binding affinity than the human 5-HT_{1B} receptor for certain adrenergic compounds caused by a difference at position 7.39 (N351 for rat and mouse, T355 for human) (*27*, *29*). Modeling of the T355N mutation into the binding pocket of human

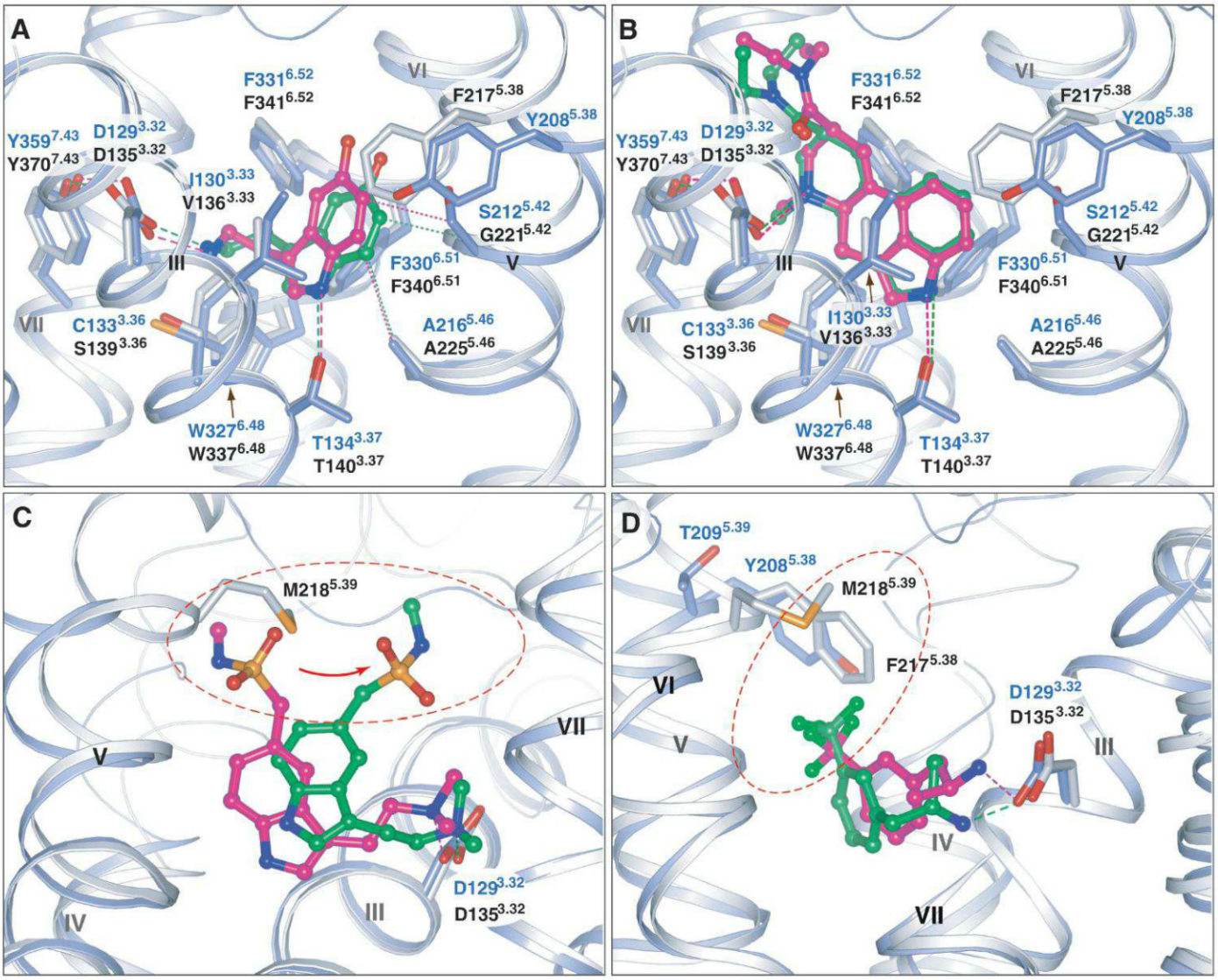


Fig. 3. Docking of the promiscuous (5-HT and LSD) and selective (sumatriptan and norfenfluoramine) ligands into the binding pockets of the 5-HT_{1B} and 5-HT_{2B} receptor structures. Docking of 5-HT (**A**), LSD (**B**), sumatriptan (**C**), and norfenfluramine (**D**) into the ligand binding pockets of 5-HT_{1B} (receptor colored light blue; ligands colored magenta) and 5-HT_{2B} (receptor colored white; ligands colored green) receptors. In (A) and (B), the polar interactions between $Y^{7.43}$ and $D^{3.32}$, and between $D^{3.32}$ and the ligands and $T^{3.37}$ and the ligands are shown as dashed lines. In (A), the nonpolar interactions between the indole ring of 5-HT and residues at positions 5.42 and 5.46 are shown as dotted lines. In (C), steric hindrance from $M218^{5.39}$ forced a reorientation of the sulfonamide group and a shift of the indole core structure of sumatriptan when docked into the 5-HT_{2B} receptor. In (D), in the 5-HT_{2B} receptor, $F217^{5.38}$ and $M218^{5.39}$ form closed contacts with the trifluoromethyl group of the ligand, which are absent in the 5-HT_{1B} receptor.

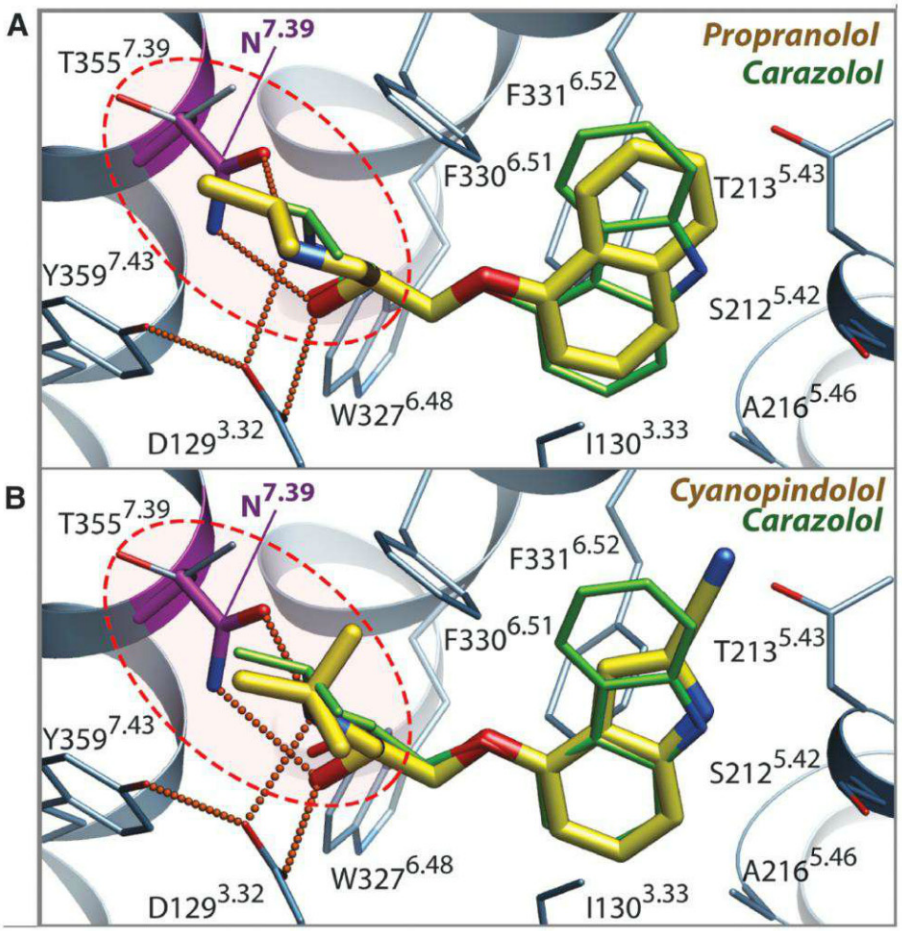


Fig. 4. Structural basis for differences in the pharmacological properties between human and rodent 5-HT$_{1B}$ receptors. The high-affinity β-adrenergic antagonists propranolol (**A**) and cyanopindolol (**B**), both shown in yellow-colored carbons, are docked in the model based on the human 5-HT$_{1B}$/ERG structure with T$^{7.39}$ mutated to N, as found in 5-HT$_{1B}$ rat and mouse orthologs. The N$^{7.39}$ side chain (magenta carbons) remained flexible in the docking procedure. The hydrogen bond network involving N$^{7.39}$, D$^{3.32}$, and Y$^{7.43}$ and propanolamine moieties of the ligands is shown as orange dots. Carazolol (green carbons) from the superimposed β$_2$-adrenergic receptor structure (PDB ID: 2RH1) is shown for comparison.

5-HT$_{1B}$ receptor revealed that N355$^{7.39}$, together with D129$^{3.32}$ and Y359$^{7.43}$, form a polar interaction network that anchors the propanolamine moiety of adrenergic antagonists, propranolol and cyanopindolol, thereby mimicking the recognition modes observed in both the β$_1$ and β$_2$ adrenergic receptors (Fig. 4). Hydrophobic interactions of I130$^{3.33}$, W327$^{6.48}$, F330$^{6.51}$, and F331$^{6.52}$ with the aromatic rings of propranolol and cyanopindolol are largely conserved between 5-HT$_{1B}$ and β-adrenergic receptors (*30*, *31*), allowing high-affinity binding of these antagonists in the human 5-HT$_{1B}$ T355$^{7.39}$N and rodent 5-HT$_{1B}$ receptors. These findings thereby provide a structural explanation for the pharmacological differences between these rodent and human GPCRs.

In conclusion, comparative analysis of the 5-HT$_{1B}$ and 5-HT$_{2B}$ receptor structures and functions, together with specific mutagenesis studies, provided a comprehensive framework for understanding ligand promiscuity and selectivity, thereby aiding development of safer and more effective medications that target the GPCR superfamily.

References and Notes

1. M. Berger, J. A. Gray, B. L. Roth, *Annu. Rev. Med.* **60**, 355 (2009).
2. D. Hoyer *et al.*, *Pharmacol. Rev.* **46**, 157 (1994).
3. B. L. Roth, S. M. Hanizavareh, A. E. Blum, *Psychopharmacology* **174**, 17 (2004).
4. B. L. Roth, D. J. Sheffler, W. K. Kroeze, *Nat. Rev. Drug Discov.* **3**, 353 (2004).
5. W. K. Kroeze, K. Kristiansen, B. L. Roth, *Curr. Top. Med. Chem.* **2**, 507 (2002).
6. M. M. Goldenberg, *Pharm. Ther.* **37**, 499 (2012).
7. B. M. Smith *et al.*, *J. Med. Chem.* **51**, 305 (2008).
8. J. Besnard *et al.*, *Nature* **492**, 215 (2012).
9. B. L. Roth, *N. Engl. J. Med.* **356**, 6 (2007).
10. C. M. Villalón, D. Centurión, L. F. Valdivia, P. de Vries, P. R. Saxena, *Curr. Vasc. Pharmacol.* **1**, 71 (2003).
11. Affected drugs include the antimigraine drugs ERG, DHE, and methysergide (usage limited); the appetite suppressant drugs fenfluramine, dex-fenfluramine, and aminorex (all withdrawn); and some drugs used for treating Parkinson's disease: pergolide and cabergoline (both withdrawn).
12. R. B. Rothman, M. H. Baumann, *Pharmacol. Biochem. Behav.* **71**, 825 (2002).
13. D. Wacker *et al.*, *Science* **340**, 615 (2013).
14. Materials and methods are available as supplementary materials on *Science* Online.
15. Superscripts refer to the Ballesteros-Weinstein numbering, in which the most conserved among class A GPCR residues in each transmembrane helix are designated x.50, where x is the helix number.
16. J. A. Ballesteros, H. Weinstein, *Methods Neurosci.* **25**, 366 (1995).
17. Residues Y177^{ECL2}, W422$^{7.35}$, and T423$^{7.36}$ in muscarinic acetylcholine receptor M2 are reported to be involved in mediating the action of allosteric modulators. Corresponding residues in the 5-HT$_{1B}$ receptor—V200^{ECL2}, F351$^{7.35}$, and D352$^{7.36}$—are part of the extended binding site.
18. K. Haga *et al.*, *Nature* **482**, 547 (2012).
19. V. Katritch, V. Cherezov, R. C. Stevens, *Trends Pharmacol. Sci.* **33**, 17 (2012).
20. M. R. Braden, D. E. Nichols, *Mol. Pharmacol.* **72**, 1200 (2007).
21. D. A. Shapiro, K. Kristiansen, W. K. Kroeze, B. L. Roth, *Mol. Pharmacol.* **58**, 877 (2000).
22. S. Mehrotra *et al.*, *Naunyn Schmiedebergs Arch. Pharmacol.* **378**, 371 (2008).
23. H. M. Connolly *et al.*, *N. Engl. J. Med.* **337**, 581 (1997).
24. R. B. Rothman, *Ann. Intern. Med.* **136**, 779 (2002).
25. R. B. Rothman *et al.*, *Circulation* **102**, 2836 (2000).
26. V. Setola, M. Dukat, R. A. Glennon, B. L. Roth, *Mol. Pharmacol.* **68**, 20 (2005).
27. N. Adham *et al.*, *Neuropharmacology* **33**, 387 (1994).
28. D. L. Nelson, V. L. Lucaites, J. E. Audia, J. S. Nissen, D. B. Wainscott, *J. Pharmacol. Exp. Ther.* **265**, 1272 (1993).
29. D. Oksenberg *et al.*, *Nature* **360**, 161 (1992).
30. V. Cherezov *et al.*, *Science* **318**, 1258 (2007).
31. T. Warne *et al.*, *Nature* **454**, 486 (2008).

Acknowledgments: This work was supported by National Institute of General Medical Sciences (NIGMS) PSI:Biology grant U54 GM094618 for biological studies and structure production (target GPCR-118) (V.K., V.C., and R.C.S.); NIH Common Fund in Structural Biology grant P50 GM073197 for technology development (V.C. and R.C.S.); the Jay and Betty Van Andel Foundation, R01 DK071662, Ministry of Science and Technology (China) grants 2012ZX09301001-005 and 2012CB910403, and Amway (China) (H.E.X.); R01 MH61887, U19 MH82441, R01 DA27170, and the National Institute of Mental Health Psychoactive Drug Screening Program (X.-P.H., E.V., J.D.M., and B.L.R.); and the Michael Hooker Chair of Pharmacology (B.L.R.). D.W. is supported by a Boehringer Ingelheim Fonds Ph.D. Fellowship. We thank J. Velasquez for help on molecular biology; T. Trinh, K. Allin, and M. Chu for help on baculovirus expression; I. Wilson for careful review and scientific feedback on the manuscript; K. Kadyshevskaya for assistance with figure preparation; A. Walker for assistance with manuscript preparation; J. Smith, R. Fischetti, and N. Sanishvili for assistance in development and use of the minibeam and beamtime at GM/CA-CAT beamline 23-ID at the Advanced Photon Source (APS), which is supported by National Cancer Institute grant Y1-CO-1020 and NIGMS grant Y1-GM-1104; and Z. Wawrzak and J. S. Brunzelle for assistance in data collection at the beamlines of sector 21 (Life Sciences Collaborative Access Team), which is funded in part by the Michigan Economic Development Corporation and Michigan Technology Tri-Corridor Grant 085P1000817. Use of the APS was supported by the Office of Science of the U.S. Department of Energy. The coordinates and the structure factors have been deposited in the Protein Data Bank (PDB) under accession codes 4IAR (5-HT$_{1B}$/ERG) and 4IAQ (5-HT$_{1B}$/DHE). R.C.S. is a founder and paid consultant for Receptos, a GPCR structure–based drug discovery company.

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1232807/DC1
Materials and Methods
Figs. S1 to S9
Tables S1 to S8
References (*32*–*53*)

15 November 2012; accepted 27 February 2013
Published online 21 March 2013;
10.1126/science.1232807

Structural Features for Functional Selectivity at Serotonin Receptors

Daniel Wacker,[1] Chong Wang,[1] Vsevolod Katritch,[1] Gye Won Han,[1] Xi-Ping Huang,[2] Eyal Vardy,[2] John D. McCorvy,[2] Yi Jiang,[1,3] Meihua Chu,[1] Fai Yiu Siu,[1] Wei Liu,[1] H. Eric Xu,[3,4] Vadim Cherezov,[1] Bryan L. Roth,[2]* Raymond C. Stevens[1]*

Drugs active at G protein–coupled receptors (GPCRs) can differentially modulate either canonical or noncanonical signaling pathways via a phenomenon known as functional selectivity or biased signaling. We report biochemical studies showing that the hallucinogen lysergic acid diethylamide, its precursor ergotamine (ERG), and related ergolines display strong functional selectivity for β-arrestin signaling at the 5-HT_{2B} 5-hydroxytryptamine (5-HT) receptor, whereas they are relatively unbiased at the 5-HT_{1B} receptor. To investigate the structural basis for biased signaling, we determined the crystal structure of the human 5-HT_{2B} receptor bound to ERG and compared it with the 5-HT_{1B}/ERG structure. Given the relatively poor understanding of GPCR structure and function to date, insight into different GPCR signaling pathways is important to better understand both adverse and favorable therapeutic activities.

Apart from canonical G protein–mediated signaling, G protein–coupled receptors (GPCRs) also activate noncanonical G protein–independent pathways, frequently mediated by β-arrestins (*1*, *2*). So-called "biased" GPCR agonists differentially activate signaling pathways with distinct efficacies and potencies as compared with unbiased agonists that activate both pathways equally (*3*). This preferential activation of one pathway over the other has been termed "functional selectivity" or "signaling bias" (*2*–*5*). Depending on the receptor, biased signaling patterns are key for mediating inflammation (*6*), apoptosis (*7*), and many other processes (*2*). Biased ligands have been proposed to stabilize receptor conformations that are distinct from those induced by unbiased ligands and selectively change the propensity of GPCR coupling to either G proteins or β-arrestins (*2*).

Agonist-induced changes in trigger motifs of GPCRs (*8*) near the binding pocket facilitate large-scale helical movements that are accompanied by rearrangements in highly conserved residues called "microswitches" (*9*) that prime GPCRs for subsequent G protein binding and activation (*10*). The structural features of a signaling-biased receptor state remain elusive, and although complexes of two β-arrestin–biased ligands with the β_1-adrenergic receptor (β_1AR) have recently been solved (*11*), they did not reveal activation-related changes in the receptor.

To elucidate molecular and structural details of biased signaling, we characterized G protein– and β-arrestin–mediated signaling at G protein–coupled serotonin [5-hydroxytryptamine (5-HT)] receptors with several representative ergolines, such as lysergic acid diethylamide (LSD) and ergotamine (ERG). Additionally, we solved the crystal structure of the 5-HT_{2B} receptor in complex with ERG, which was identified as a highly biased agonist for the 5-HT_{2B} receptor (*12*).

To investigate potential differences of ergoline signaling at 5-HT receptors, we examined three prototypical serotonin receptors that interact with distinct G proteins. The 5-HT_{1B} receptor inhibits cyclic adenosine monophosphate (cAMP) production through G_i, the 5-HT_{2B} receptor mediates phospholipase C activation through G_q, and the 5-HT_{7A} receptor stimulates cAMP production through G_s (*13*). We compared G protein– and β-arrestin–mediated signaling at cloned human 5-HT_{1B} and 5-HT_{2B} receptors and G protein–mediated signaling at 5-HT_{7A} receptors stimulated by selective and nonselective ligands in human embryonic kidney (HEK) 293 cells (Fig. 1 and table S1) (*14*).

LSD and, especially, ERG displayed bias for β-arrestin signaling at 5-HT_{2B} (bias factors 101 and 228, respectively) (Fig. 1D), minimal bias at 5-HT_{1B} (bias factors 5 and 25, respectively) (Fig. 1D), and G protein antagonism at 5-HT_{7A} receptors (Fig. 1B and table S1). We also found significant β-arrestin signaling bias for other ergolines—such as dihydroergotamine (DHE), methylergonovine (MTE), pergolide (PER), and cabergoline (CAB)—at the 5-HT_{2B} receptor, whereas all other evaluated compounds showed no significant bias (Fig. 1D). ERG and DHE, both of which contain a large tripeptide moiety substitution at the amide scaffold, displayed more extreme signaling bias at the 5-HT_{2B} receptor compared with LSD.

To investigate the molecular details responsible for biased signaling, we crystallized an engineered 5-HT_{2B} receptor construct in complex with ERG, solved its structure at 2.7 Å (fig. S1 and S2 and table S3), and compared it with the structure of 5-HT_{1B}/ERG reported in the companion manuscript (*15*), as well as to other known unbiased active-state GPCR structures. Residues $P^{5.50}$, $I^{3.40}$, and $F^{6.44}$ (*16*, *17*), the "P-I-F" motif (P, Pro; I, Ile; F, Phe), form an interface between helices V, III, and VI near the base of the ligand binding pocket in β_2AR and many other aminergic receptors, including all 5-HT GPCRs. In the active-state structures of β_2AR (*8*, *18*), a chain of conformational rearrangements occurs in the P-I-F residues, in which an inward shift of helix V residue $P211^{5.50}$ is coupled with: (i) a rotamer switch in $I121^{3.40}$, (ii) a large movement of the $F282^{6.44}$ side chain, and (iii) a corresponding rotation of helix VI on the cytoplasmic side (*8*). The 5-HT_{1B} and 5-HT_{2B} receptor structures display two different conformations of the P-I-F motif (Fig. 2). For the 5-HT_{1B} receptor, we observe that the P-I-F configuration is essentially identical to that of the active-state of β_2AR [β_2AR-R*, Protein Data Bank identification number (PDB ID): 3SN6 (*18*)] (Fig. 2B). Whereas the 5-HT_{2B} receptor adopts a similar active-like conformation of $P229^{5.50}$ and $I143^{3.40}$, the side-chain conformation of $F333^{6.44}$ was similar to that observed in the inactive β_2AR [β_2AR-R, PDB ID: 2RH1 (*19*)] (Fig. 2C and fig. S2B). The P-I-F motif, therefore, appears to be in an active-like conformation in the 5-HT_{1B} structure, but only in an intermediate active conformation in the 5-HT_{2B} receptor structure.

At the cytoplasmic side of the receptors, GPCR activation is generally characterized by the displacement of helices V, VI, and VII (*10*). The magnitude of the helical motions depends on the activation-state; for example, the outward displacement of the intracellular tip of helix VI ranges between 3 and 14 Å, whereas helix VII shifts between 3 to 5 Å toward the receptor core (*10*, *18*–*22*). These concerted rearrangements induce an opening of the helical bundle at the cytoplasmic side, which facilitates the binding and subsequent activation of G proteins. Analysis of the 5-HT_{1B} and 5-HT_{2B} receptor structures shows that the conformation in the intracellular half of the helical bundle is notably shifted toward that seen in active-state GPCR structures (Fig. 3 and figs. S3 and S4) and is distinct from those of inactive-state GPCR structures (Fig. 3 and figs. S5 and S6). In the 5-HT_{1B} and 5-HT_{2B} receptor structures, the helix VI intracellular part is located at least 2 to 4 Å further away from the receptor core than in the inactive-state structures of other aminergic GPCRs (fig. S5). This conformation of helix VI is close to the one observed in active-state structures of the A_{2A} adenosine receptor (A_{2A}AR) (fig. S3) and rhodopsin (Rho) (fig. S4), though the outward shift of the helix is smaller in magnitude compared with that of the G protein–bound β_2AR (Fig. 3, A and C). The only difference that we observed in helix VI between 5-HT receptor subtypes was a small clockwise rotation in the 5-HT_{1B} receptor toward the active state that is

[1]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA. [2]National Institute of Mental Health Psychoactive Drug Screening Program, Department of Pharmacology and Division of Chemical Biology and Medicinal Chemistry, University of North Carolina Chapel Hill Medical School, Chapel Hill, NC 27599, USA. [3]Van Andel Research Institute/Shanghai Institute of Materia Medica Center, CAS-Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China. [4]Laboratory of Structural Sciences, Van Andel Research Institute, 333 Bostwick Ave NE, Grand Rapids, MI 49503, USA.

*Corresponding author. E-mail: bryan_roth@med.unc.edu (B.L.R.); stevens@scripps.edu (R.C.S.)

absent in the 5-HT_{2B} structure (Figs. 2, B and C, and 3, A and C). Helix VII in both 5-HT receptor structures also displays intermediate active states when compared with β_2AR (Fig. 3, B and D), where it is shifted toward the receptor core as compared with inactive-state structures of other aminergic GPCRs (fig. S6). Whereas the 5-HT_{2B}/ERG receptor structure shows less pronounced active-like changes in helix VI, helix VII appears to be in a more active conformation than it is in the 5-HT_{1B}/ERG receptor structure (Fig. 3, B and D).

Another important aspect of GPCR activation is the rearrangement of side chains in highly conserved motifs D(E)/RY (helix III) and NPxxY (helix VII) (D, Asp; E, Glu; R, Arg; Y, Tyr; N, Asn; x, any amino acid), which are referred to as microswitches (*9*). Thus, the D(E)/RY motif

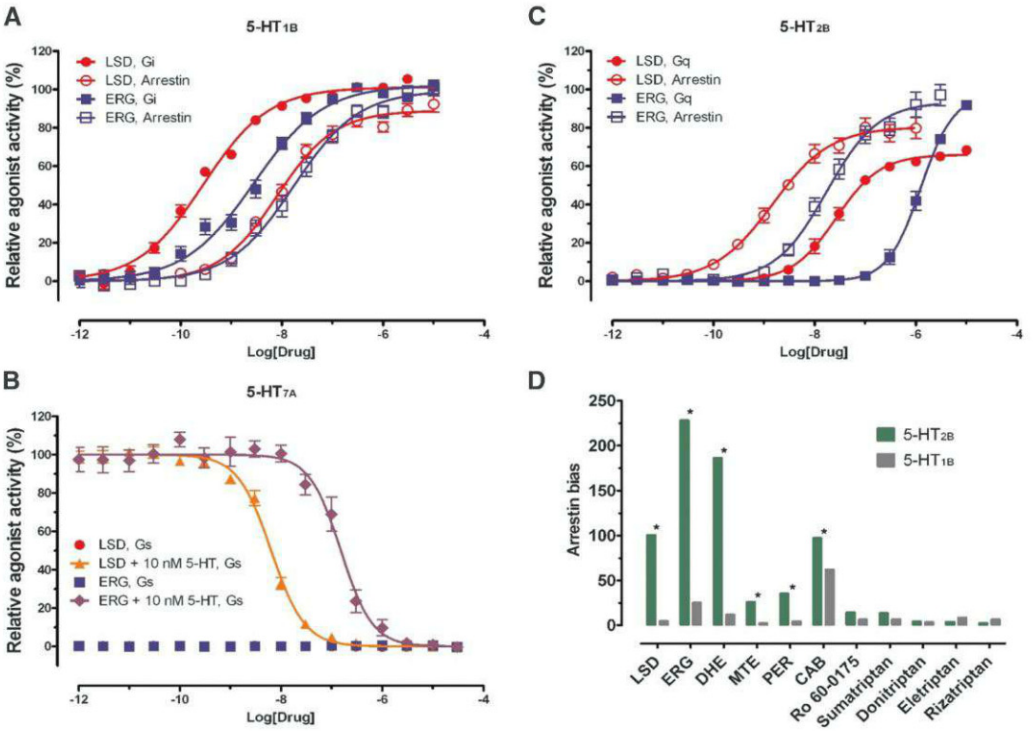


Fig. 1. Distinct signaling properties of LSD and ERG at 5-HT_{1B}, 5-HT_{7A}, and 5-HT_{2B} receptors. We used luminescence-based assays to measure 5-HT_{1B} receptor–mediated G_i activation and cAMP production, fluorescence-based calcium mobilization assays to measure 5-HT_{2B} receptor–mediated G_q activation, and β-arrestin translocation-dependent luciferase reporter assays to measure 5-HT_{1B} and 5-HT_{2B} receptor–mediated β-arrestin recruitment, all in HEK293 derived cells. (**A**) Normalized concentration-response studies for LSD and ERG at human cloned 5-HT_{1B} receptor–mediated activation of G_i and noncanonical (arrestin) signaling. (**B**) Normalized concentration-response studies for LSD and ERG at human cloned 5-HT_{7A} receptor–mediated activation of G_s signaling in the presence and absence of 5-HT. The solid red circles for LSD-G_s signals are superimposed by the solid blue squares for ERG-G_s signals and, thus, not visible. (**C**) Normalized concentration-response studies for LSD and ERG at human cloned 5-HT_{2B} receptor–mediated activation of G_q and noncanonical (arrestin) signaling. (**D**) Mean β-arrestin bias factors were calculated for serotonergic agonists at 5-HT_{2B} and 5-HT_{1B} receptors. Concentration-responses curves were fit to the Black and Leff operational model to obtain transduction coefficients [Log(τ/K_A)] (where τ is agonist efficacy and K_A is the equilibrium dissociation constant) for each ligand at each corresponding pathway. The ΔLog(τ/K_A) was then calculated with 5-HT as a reference agonist for each pathway, and the $\Delta\Delta$Log(τ/K_A) was calculated between two pathways for each ligand. The bias factor is unitless and defined as $10^{\Delta\Delta \mathrm{Log}\ \tau/K_A}$ (*28*). Compounds with values close to one represent unbiased agonists, whereas compounds with large numerical values, typically >100, represent extremely biased agonists. *P < 0.0001 via two-way analysis of variance comparing 5-HT_{2B} versus 5-HT_{1B} bias factors; n = 3 to 6 separate experiments. ERG, DHE, and, to a lesser extent, LSD, MTE, and PER show strong β-arrestin bias at the 5-HT_{2B} receptor, but not the 5-HT_{1B} receptor. Error bars in (A) to (C) denote SEM from a minimum of three assays.

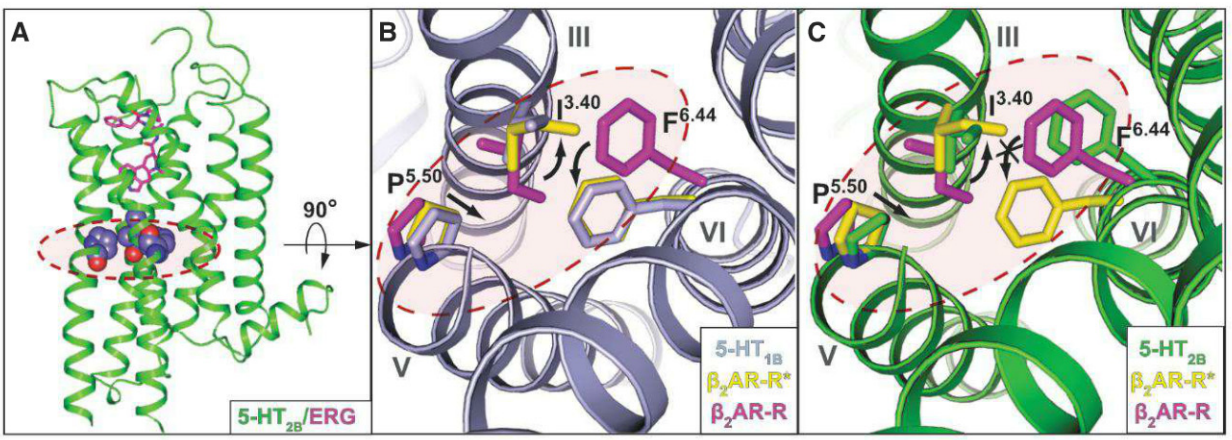


Fig. 2. Trigger motif P-I-F displays active and intermediate active states for 5-HT_{1B} and 5-HT_{2B} receptors, respectively. Residues of the P-I-F motif are highlighted in a dashed red circle in each panel. (**A**) Overall architecture of the 5-HT_{2B} receptor (green) bound to ERG (magenta); residues of the $P^{5.50}$ $I^{3.40}$ $F^{6.44}$ motif are illustrated in space-filling representation. (**B**) Alignment between 5-HT_{1B} receptor (gray), β_2AR-R (magenta; PDB ID: 2RH1), and β_2AR-R*(yellow; PDB ID: 3SN6) indicates an activated P-I-F motif in the 5-HT_{1B}/ERG structure. Black arrows denote likely rearrangements upon 5-HT_{1B} receptor activation, according to analysis of β_2AR. (**C**) Alignment between 5-HT_{2B} receptor (green), β_2AR-R (magenta), and β_2AR-R* (yellow) suggests an intermediate active state of the P-I-F motif in the 5-HT_{2B}/ERG structure, with $F^{6.44}$ in an inactive conformation. Black arrows indicate likely rearrangements upon 5-HT_{2B} receptor activation, according to analysis of β_2AR.

of all inactive-state and most active-state GPCR structures shows an intact salt bridge between the side chains of D(E)$^{3.49}$ and R$^{3.50}$. Importantly, this salt bridge is broken only in the active-state structures of β_2AR-G$\alpha\beta\gamma$ (*18*) and Opsin-GαCT (*21*), where R$^{3.50}$ interacts instead with the G protein and Gα peptide, respectively. The salt bridge is preserved between the side chains of R153$^{3.50}$ and D152$^{3.49}$ in the 5-HT$_{2B}$ receptor structure, but it is broken in the 5-HT$_{1B}$ receptor structure (Fig. 4, A, B, and D; and fig. S2C). In the 5-HT$_{1B}$ receptor structure, the D146$^{3.49}$ side chain forms a hydrogen bond to Y157 in intracellular loop 2, and the R147$^{3.50}$ side chain interacts with the main-chain carbonyl in a loop of the fusion protein BRIL (residue L1048; L, Leu), which is partially inserted into the G protein binding crevice (Fig. 4A). Thus, the conformation of the D(E)RY motif mimics the active state of β_2AR in the 5-HT$_{1B}$ structure but resembles the inactive state in the 5-HT$_{2B}$ structure (Fig. 4, A and C; and fig. S7).

The highly conserved NPxxY motif at the cytoplasmic end of helix VII is another key microswitch of GPCR activation (*9*). Upon GPCR activation, the intracellular end of helix VII moves toward the receptor core, and a rotation of Y$^{7.53}$ around the helical axis moves the side chain further into the seven-transmembrane (TM) bundle (*10*). Both 5-HT receptor structures show active-state conformations of the NPxxY motif when compared with β_2AR, A_{2A}AR, and Rho (Fig. 4, D and E; and figs. S2D and S7), with more pronounced activation features in the 5-HT$_{2B}$ receptor.

Our analysis indicates that the 5-HT$_{1B}$/ERG structure has most of the attributes of a classical agonist-induced, active-like state, consistent with our biochemical findings that ERG is a comparatively unbiased agonist at the 5-HT$_{1B}$ receptor. In contrast, the 5-HT$_{2B}$/ERG structure exhibits conformational characteristics of both the active and inactive states. The structure of β_2AR-G$\alpha\beta\gamma$ complex, along with recent nuclear magnetic resonance and fluorescence studies of β_2AR, implicates helix VI predominantly in G protein signaling, whereas conformational changes in helix VII are associated with enhanced β-arrestin signaling (*18*, *23*, *24*). Thus, an active-like state in the helix VII conformation of the 5-HT$_{2B}$ receptor, but only partial changes in helix VI, mirrors the strong β-arrestin bias of ERG at 5-HT$_{2B}$ receptors observed in pharmacological assays.

A likely structural explanation for the distinct conformational features and biased pharmacology of ERG between 5-HT$_{1B}$ and 5-HT$_{2B}$ receptors can be found in the region of the extracellular loop 2 (ECL2) junction with helix V. In the 5-HT$_{2B}$ receptor structure, E212-R213-F214 forms an additional helical turn stabilized by a structured water molecule at the extracellular tip of helix V (Fig. 5). As a result, the segment of ECL2 that connect helices III and V via the conserved disulfide bond is shortened in the 5-HT$_{2B}$ receptor, inducing an inward shift and creating a conformational constraint on the position of the extracellular tip of helix V. Because of these rearrangements, ERG forms additional hydrophobic contacts with M218$^{5.39}$, L347$^{6.58}$, V348$^{6.59}$, L362$^{7.35}$, and K211^{ECL2} (M, Met; V, Val; K, Lys) (fig. S8A) and stabilizes a closer distance between helices V and VI, which form hydrophobic contacts between L219$^{5.40}$, V348$^{6.59}$, and L349$^{6.60}$, along with a hydrogen bond between S222$^{5.43}$ and N344$^{6.55}$ (S, Ser) (fig. S8B). These extensive ligand-mediated interactions between helices V and VI in the 5-HT$_{2B}$/ERG complex may be inferred to prevent rearrangements in helix VI and the corresponding rotation of F$^{6.44}$ (*8*), observed in the 5-HT$_{1B}$ receptor and structures of other active-state GPCRs. The strengthened interactions of helix V and VI through ligand-mediated hydrogen bonds to both helices have also been linked to inhibition of G protein signaling at β_2AR by the most efficacious inverse agonist (*25*).

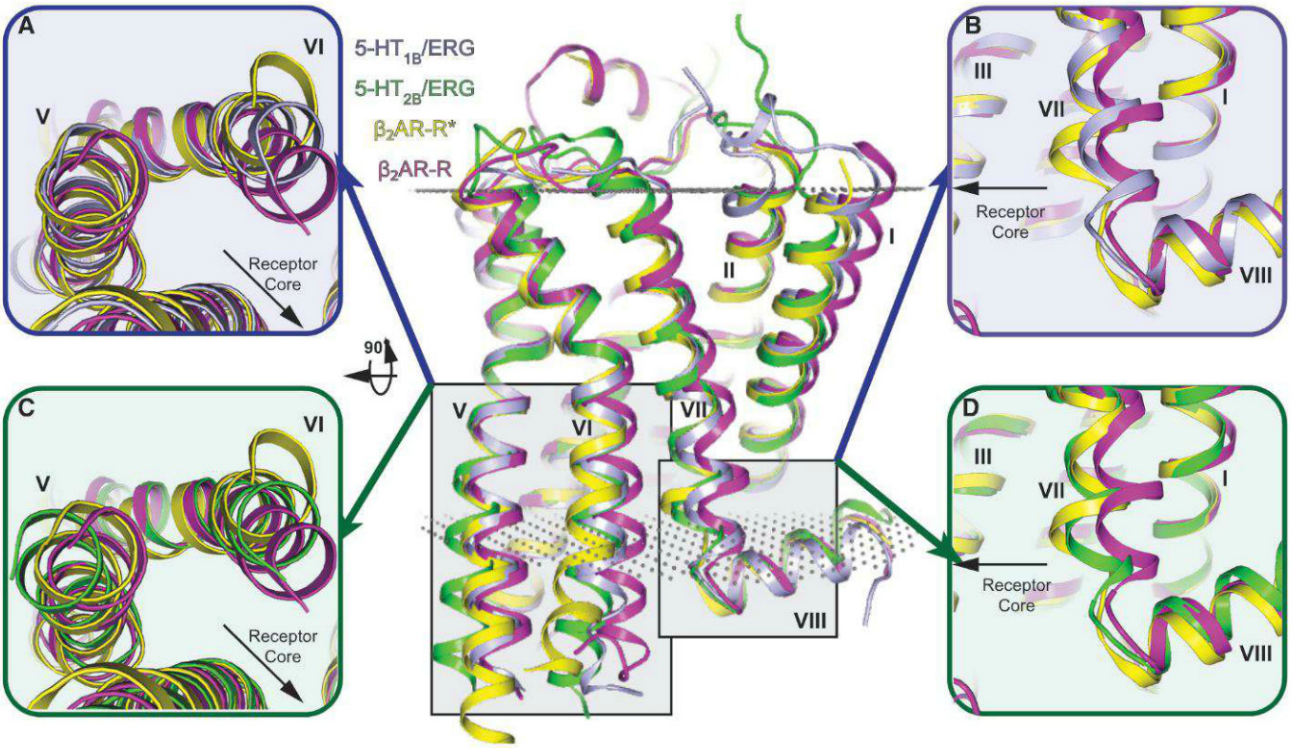


Fig. 3. Structural alignment with β_2AR-R and β_2AR-R* reveals distinct active-state seven-TM conformations of the 5-HT$_{1B}$ and 5-HT$_{2B}$ receptor structures. All structures were analyzed based on the last membrane-embedded residue to minimize the effect of G proteins and fusion partners on the relative helix positions (see supplementary materials). The center panel shows the overall seven-TM configuration of β_2AR-R (magenta; PDB ID: 2RH1), β_2AR-R*(yellow; PDB ID: 3SN6), 5-HT$_{1B}$ (gray), and 5-HT$_{2B}$ (green) receptors aligned through helices I to IV. Membrane boundaries are indicated by gray dots according to the Orientations of Proteins in Membranes database (*29*). (**A** and **C**) Intracellular view of helices V and VI in β_2AR-R, β_2AR-R* compared with (A) the 5-HT$_{1B}$ receptor or (C) the 5-HT$_{2B}$ receptor. Residues on the intracellular side of the membrane have been removed for better comparison of ligand-induced helical rearrangements (see main text). Helix VI of the 5-HT$_{1B}$ and 5-HT$_{2B}$ receptors is in an intermediate active state compared with β_2AR. (**B** and **D**) Side view of helices VII and VIII in β_2AR-R, β_2AR-R* and the (B) 5-HT$_{1B}$ receptor or (D) 5-HT$_{2B}$ receptor. Helix VII of the 5-HT$_{1B}$ and 5-HT$_{2B}$ receptors is in an intermediate active state compared with β_2AR, with more pronounced activation features for the 5-HT$_{2B}$ receptor.

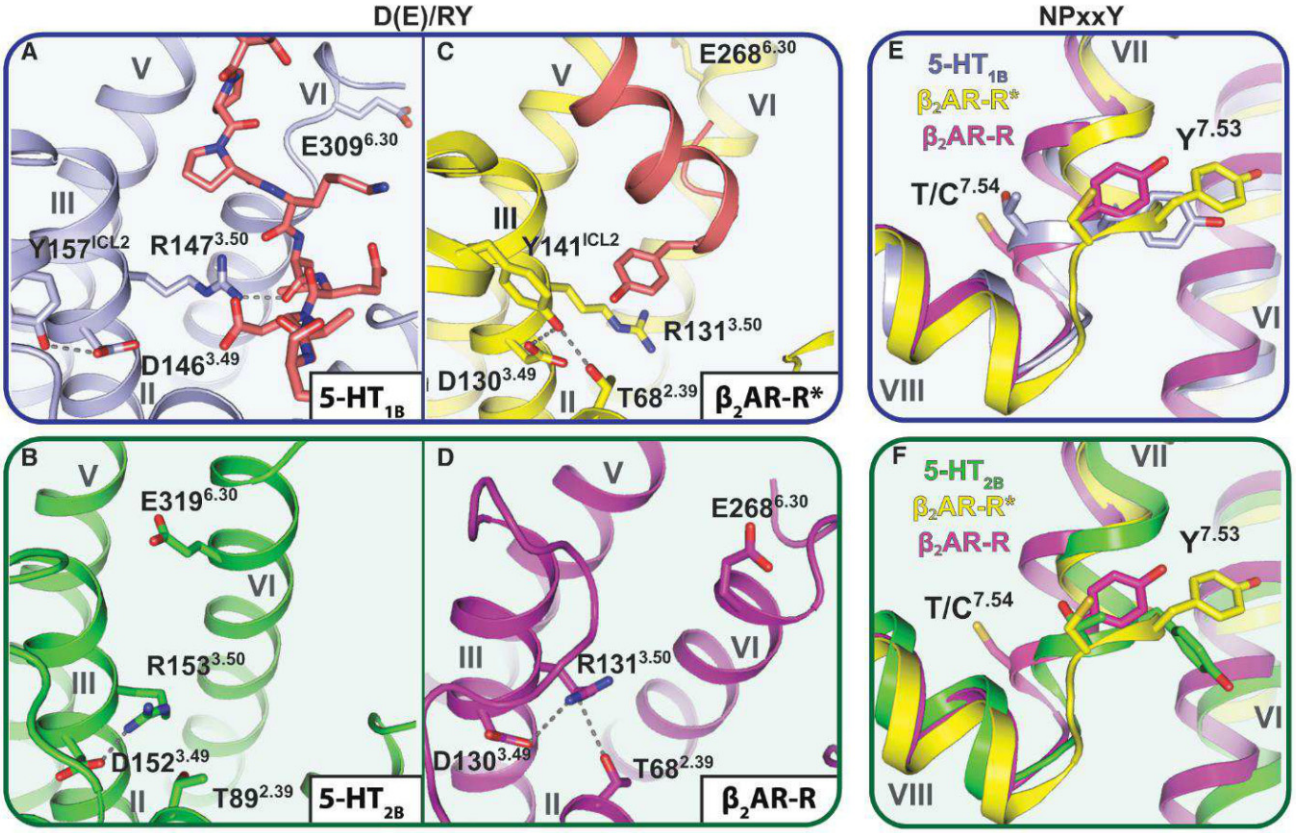


Fig. 4. Activation state of the D(E)/RY and NPxxY motifs in 5-HT$_{1B}$ and 5-HT$_{2B}$ receptors compared with β$_2$AR-R and β$_2$AR-R*. Configuration of the D(E)/RY motif in (**A**) the 5-HT$_{1B}$ receptor (gray), (**B**) the 5-HT$_{2B}$ receptor (green), (**C**) β$_2$AR-R* (yellow; PDB ID: 3SN6), and (**D**) β$_2$AR-R (magenta; PDB ID: 2RH1). Residues of the G protein in β$_2$AR-R* (C) and the G protein–mimicking BRIL loop (A) are highlighted in orange. The conformation of the D(E)/RY motif in the 5-HT$_{1B}$ receptor is similar to that observed in β$_2$AR-R*, whereas the configuration of the 5-HT$_{2B}$ receptor compares to that of β$_2$AR-R. (**E** and **F**) Conformational states of Y$^{7.53}$ of the NPxxY motif and the proceeding residue 7.54 in β$_2$AR-R, β$_2$AR-R*, and (E) the 5-HT$_{1B}$ receptor or (F) 5-HT$_{2B}$ receptor. When compared with β$_2$AR, the conformation of the NPxxY motif in the 5-HT$_{1B}$ receptor is in an intermediate active state, whereas the configuration of the 5-HT$_{2B}$ receptor is similar to β$_2$AR-R*. T, Thr; C, Cys.

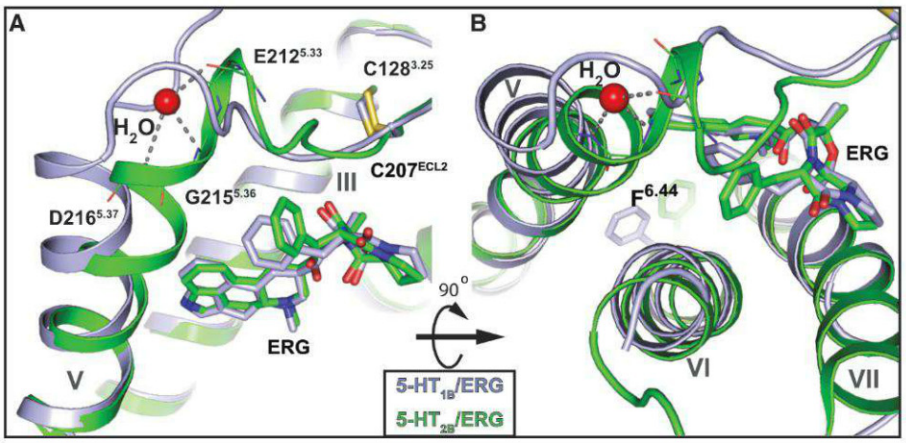


Fig. 5. Structural differences in extracellular configuration of helix V between 5-HT$_{1B}$ and 5-HT$_{2B}$ receptors likely explain β-arrestin functional selectivity at the 5-HT$_{2B}$ receptor. Side view (**A**) and top view (**B**) of 5-HT$_{1B}$ (gray) and 5-HT$_{2B}$ (green) receptors show a kink in the extracellular end of helix V in the 5-HT$_{2B}$ receptor. A water molecule (red sphere) was found to stabilize the kink through hydrogen bonds (gray dashed lines) with the E212$^{5.33}$ main-chain carbonyl oxygen and the main-chain nitrogens of D216$^{5.37}$ and G215$^{5.36}$ (G, Gly). Main-chain atoms of both residues are shown as lines.

Ergolines predominantly signal through β-arrestin pathways at 5-HT$_{2B}$ receptors, whereas signaling at 5-HT$_{1B}$ receptors appears nonbiased. The differential signaling patterns are mirrored in the crystal structures, which show features of an intermediate active state for the 5-HT$_{1B}$ receptor and a β-arrestin–biased activation state for the 5-HT$_{2B}$ receptor. We propose a mechanism by which ERG stabilizes a conformation of the 5-HT$_{2B}$ receptor that selectively interferes with G protein signaling (fig. S9). The tripeptide moiety of ERG appears to interfere with G protein signaling at the 5-HT$_{2B}$ receptor, as ERG exhibits strongly increased β-arrestin signaling bias compared with LSD and MTE. Because both therapeutic (*26*) and adverse (*27*) drug effects have been associated with β-arrestin recruitment by GPCRs, identifying the features of biased GPCR states by known and yet to be discovered intracellular signaling proteins may facilitate the development of safer and more effective therapeutics with selective signaling profiles.

References and Notes

1. L. M. Luttrell *et al.*, *Science* **283**, 655 (1999).
2. E. Reiter, S. Ahn, A. K. Shukla, R. J. Lefkowitz, *Annu. Rev. Pharmacol. Toxicol.* **52**, 179 (2012).
3. T. Kenakin, *J. Pharmacol. Exp. Ther.* **336**, 296 (2011).
4. J. A. Allen, B. L. Roth, *Annu. Rev. Pharmacol. Toxicol.* **51**, 117 (2011).
5. J. D. Urban *et al.*, *J. Pharmacol. Exp. Ther.* **320**, 1 (2007).
6. H. Gao *et al.*, *Mol. Cell* **14**, 303 (2004).
7. S. Ahn, J. Kim, M. R. Hara, X. R. Ren, R. J. Lefkowitz, *J. Biol. Chem.* **284**, 8855 (2009).

8. S. G. Rasmussen *et al.*, *Nature* **469**, 175 (2011).
9. R. Nygaard, T. M. Frimurer, B. Holst, M. M. Rosenkilde, T. W. Schwartz, *Trends Pharmacol. Sci.* **30**, 249 (2009).
10. V. Katritch, V. Cherezov, R. C. Stevens, *Annu. Rev. Pharmacol. Toxicol.* **53**, 531 (2013).
11. T. Warne, P. C. Edwards, A. G. Leslie, C. G. Tate, *Structure* **20**, 841 (2012).
12. X. P. Huang *et al.*, *Mol. Pharmacol.* **76**, 710 (2009).
13. M. Berger, J. A. Gray, B. L. Roth, *Annu. Rev. Med.* **60**, 355 (2009).
14. Materials and methods and supplementary figures and tables are available as supplementary materials on *Science* Online.
15. C. Wang *et al.*, *Science* **340**, 610 (2013).
16. Superscripts refer to the Ballesteros-Weinstein numbering, in which the most conserved among class A GPCRs residues in each TM helix are designated x.50, where x is the helix number.
17. J. A. Ballesteros, H. Weinstein, *Methods Neurosci.* **25**, 366 (1995).
18. S. G. Rasmussen *et al.*, *Nature* **477**, 549 (2011).
19. V. Cherezov *et al.*, *Science* **318**, 1258 (2007).
20. Y. Imamoto, M. Kataoka, F. Tokunaga, K. Palczewski, *Biochemistry* **39**, 15225 (2000).
21. P. Scheerer *et al.*, *Nature* **455**, 497 (2008).
22. F. Xu *et al.*, *Science* **332**, 322 (2011).
23. J. J. Liu, R. Horst, V. Katritch, R. C. Stevens, K. Wüthrich, *Science* **335**, 1106 (2012).
24. R. Rahmeh *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 6733 (2012).
25. D. Wacker *et al.*, *J. Am. Chem. Soc.* **132**, 11443 (2010).
26. J. A. Allen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 18488 (2011).
27. A. K. Finn, J. L. Whistler, *Neuron* **32**, 829 (2001).
28. T. Kenakin, C. Watson, V. Muniz-Medina, A. Christopoulos, S. Novick, *ACS Chem. Neurosci.* **3**, 193 (2012).
29. M. A. Lomize, I. D. Pogozheva, H. Joo, H. I. Mosberg, A. L. Lomize, *Nucleic Acids Res.* **40**, D370 (2012).

Acknowledgments: This work was supported by the National Institute of General Medical Sciences (NIGMS) Protein Structure Initiative:Biology grant U54 GM094618 for biological studies and structure production (target GPCR-4) (V.K., V.C., and R.C.S.); NIH Common Fund in Structural Biology grant P50 GM073197 for technology development (V.C. and R.C.S.); the Jay and Betty Van Andel Foundation, Amway (China), grant R01 DK071662; Ministry of Science and Technology (China) grants 2012ZX09301001-005 and 2012CB910403 (H.E.X.); grants U19 MH82441, R01 MH61887, and the National Institute of Mental Health Psychoactive Drug Screening Program (X.-P.H., E.V., and B.L.R.); and the Michael Hooker Chair of Pharmacology (B.L.R.). D.W. is supported by a Boehringer Ingelheim Fonds Ph.D. Fellowship. R.C.S. is a founder and paid consultant for Receptos, a GPCR structure–based drug discovery company. We thank J. Velasquez for help on molecular biology; T. Trinh, K. Allin, and M. Chu for help on baculovirus expression; L. N. Collins for help on initial construct selection; K. Kadyshevskaya for assistance with figure preparation; A. Walker for assistance with manuscript preparation; I. Wilson for careful review and scientific feedback on the manuscript; T. Kenakin (University of North Carolina) for helpful discussions regarding the quantification of ligand bias; J. Smith, R. Fischetti, and N. Sanishvili for assistance in development and use of the minibeam and beamtime at GM/CA-CAT beamline 23-ID at the Advanced Photon Source, which is supported by National Cancer Institute grant Y1-CO-1020 and NIGMS grant Y1-GM-1104. Use of the Advanced Photon Source was supported by the Office of Science of the U.S. Department of Energy. Coordinates and the structure factors of the 5-HT_{2B}/ERG complex have been deposited in the Protein Data Bank under the accession code 4IB4.

Supplementary Materials
www.sciencemag.org/cgi/content/full/science.1232808/DC1
Materials and Methods
Figs. S1 to S9
Tables S1 to S3
References (*30*–*42*)

15 November 2012; accepted 27 February 2013
Published online 21 March 2013;
10.1126/science.1232808

R-Loop Stabilization Represses Antisense Transcription at the *Arabidopsis FLC* Locus

Qianwen Sun,[1] Tibor Csorba,[1] Konstantina Skourti-Stathaki,[2] Nicholas J. Proudfoot,[2] Caroline Dean[1]*

Roles for long noncoding RNAs (lncRNAs) in gene expression are emerging, but regulation of the lncRNA itself is poorly understood. We have identified a homeodomain protein, AtNDX, that regulates *COOLAIR*, a set of antisense transcripts originating from the 3′ end of *Arabidopsis FLOWERING LOCUS C* (*FLC*). AtNDX associates with single-stranded DNA rather than double-stranded DNA non–sequence-specifically in vitro, and localizes to a heterochromatic region in the *COOLAIR* promoter in vivo. Single-stranded DNA was detected in vivo as part of an RNA-DNA hybrid, or R-loop, that covers the *COOLAIR* promoter. R-loop stabilization mediated by AtNDX inhibits *COOLAIR* transcription, which in turn modifies *FLC* expression. Differential stabilization of R-loops could be a general mechanism influencing gene expression in many organisms.

A major factor determining natural variation in flowering in *Arabidopsis* is quantitative variation in the expression and silencing of the floral repressor gene *FLC* (*1*, *2*). Multiple pathways regulate *FLC*, and these converge on cotranscriptional mechanisms involving antisense transcripts (named *COOLAIR*) and different chromatin pathways (*3*, *4*). One of these pathways is vernalization, when prolonged cold increases *COOLAIR* transcription and induces a Polycomb-mediated epigenetic silencing of *FLC* (*5*, *6*). Another is the autonomous pathway, which involves alternative processing of *COOLAIR* transcripts that causes gene body histone K4 demethylation and *FLC* down-regulation (*4*, *7*). Because these regulators are conserved through evolution, *COOLAIR* regulation may have the potential to inform long noncoding RNA (lncRNA) function generally (*8*, *9*).

To investigate *COOLAIR* regulation, we undertook a forward mutagenesis screen using a luciferase reporter system (*pCOOLAIR:Luc*) (fig. S1) (*3*). *eoc1* (*enhancer of COOLAIR1*) was identified and mapped to a 32,000–base pair (bp) region on chromosome 4 (Fig. 1A and fig. S2, A to D). A G-A transition was detected in the ninth exon of At4g03090 that resulted in a premature stop codon (TGG to TGA). Complementation and allelic analysis confirmed that the enhanced Luc signal in *eoc1*-1 was due to the mutation of At4g03090 gene, previously named *AtNDX* (fig. S3, A to C) (*10*).

AtNDX is an atypical and highly divergent plant homeodomain (HD)–containing protein conserved in moss, *Selaginella*, and other flowering plants (Fig. 1B and fig. S4) (*10*). Analysis of insertion mutants (*eoc1*-2, *eoc1*-4) showed that AtNDX represses endogenous *COOLAIR* expression (Fig. 1, B and C). Relative to wild-type plants, the *eoc1*-2 and *eoc1*-4 mutants flowered later (fig. S5B) and their *FLC* expression was up-regulated (fig. S5C). The defect in *COOLAIR* expression was rescued by expression of FLAG-tagged AtNDX (fig. S6). AtNDX is expressed predominantly in dividing tissues such as young leaves, root tips, flower buds, and embryos (fig. S7).

The presence of the HD domain (Fig. 1B and fig. S4) and the localization of green fluorescent protein–tagged AtNDX in the nucleus and nucleolus (fig. S7C and fig. S8) prompted us to test whether AtNDX associates with *FLC* chromatin. Chromatin immunoprecipitation (ChIP) experiments using the FLAG-tagged AtNDX lines showed that AtNDX is enriched at the *COOLAIR* promoter–*FLC* terminator region (Fig. 2A). This suggests that the effects of AtNDX on *COOLAIR* are direct. GmNDX, the homolog of AtNDX in soybean, shows DNA binding via its HD domain (*11*). However, HD proteins can also bind RNA (*12*). To test the binding properties of AtNDX, we performed electrophoretic mobility shift assay (EMSA) analysis using a glutathione *S*-transferase (GST) recombinant protein that included the HD and NDX-B domain (GST-AtNDX; Fig. 1B and fig. S9A). We did not see binding of GST-AtNDX to double-stranded DNA (dsDNA) probes (Fig. 2B and fig. S9), but GST-AtNDX bound single-stranded DNA (ssDNA) in a non–sequence-specific manner (Fig. 2B, fig. S9D, and table S1). No binding to ssRNA, dsRNA, or RNA-DNA hybrids was detected (Fig. 2B, fig. S9, D and E, and table S1).

Single-stranded DNA can be formed in vivo during transcription if nascent RNA transcripts invade the dsDNA and anneal to the template strand in the duplex, generating an RNA-DNA hybrid. A three-stranded nucleic acid structure formed by an RNA-DNA hybrid plus a displaced ssDNA strand is called an R-loop (*13*). R-loops have been considered as transcriptional

[1]Department of Cell and Developmental Biology, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, UK. [2]Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, UK.

*Corresponding author. E-mail: caroline.dean@jic.ac.uk

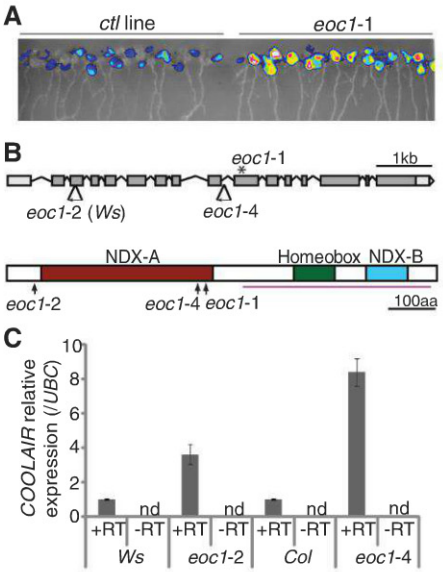


Fig. 1. The HD domain protein AtNDX is a regulator of *FLC*. (**A**) A mutant showing enhanced luciferase activity was isolated and called *eoc1* (*enhancer of COOLAIR1*). Luciferase activity is depicted with false color (blue < green < yellow < red < white). (**B**) *eoc1* maps to AtNDX, which encodes a homeodomain-containing protein with two other conserved domains, NDX-A and NDX-B (fig. S4) (*10*). *eoc1*-1 is a point mutation; *eoc1*-2 and *eoc1*-4 are T-DNA insertion alleles in *Ws* and *Col*-0 background, respectively. The pink line indicates the recombinant protein used for EMSA (fig. S9A); aa, amino acids. (**C**) Endogenous *COOLAIR* RNA is upregulated in *eoc1*; expression data are relative to *UBC21*, normalized to the wild type. Errors are SEM for three biological replicates.

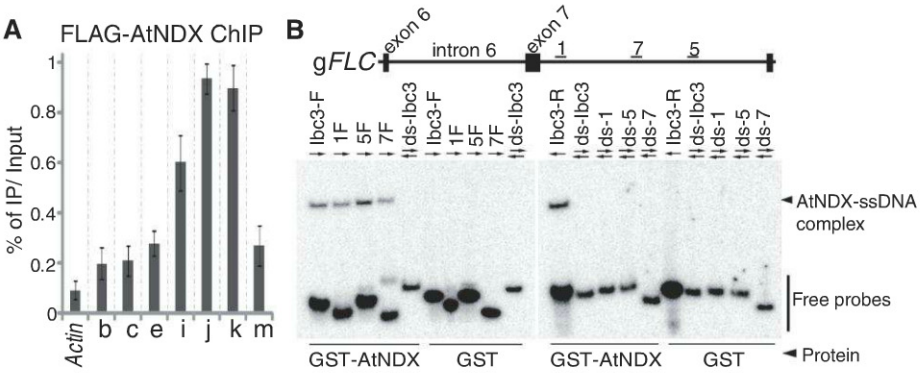


Fig. 2. AtNDX associates with the *COOLAIR* promoter in vivo and has ssDNA binding specificity in vitro. (**A**) ChIP–quantitative PCR (qPCR) analysis of AtNDX binding along *FLC*. *ACTIN2* was used as a negative control. Errors are SEM of three biological replicates; two repeats of qPCR were tested for each ChIP replicate. The regions tested are shown in Fig. 3B. (**B**) Recombinant GST-AtNDX binds to ssDNA (one arrow) but not dsDNA (two arrows). Sequences of *lbc3* probes were from (*11*). Probes 1, 5, and 7 from the *FLC* locus are shown in the schematic; ssDNA and dsDNA probes were purified and labeled (fig. S9C). GST alone was used as a negative control.

by-products, but recent data suggest that R-loops may have an impact on gene expression (*14*–*16*). In *Saccharomyces cerevisiae*, R-loops impair RNA polymerase II (Pol II) transcription elongation (*16*). In mammals, an RNA/DNA helicase called senataxin resolves R-loops, and this helps transcription termination and Pol II release (*15*). R-loops tend to form in GC-rich genomic regions (*17*, *18*), and recent evidence suggests that R-loop formation may maintain an unmethylated DNA state at promoters with skewed CpG islands, correlating positively with transcriptional activity in mammals (*17*). The *COOLAIR* promoter region is rich in GC nucleotides (around 60% GC; fig. S10B), is very low in DNA methylation (*19*), and has low nucleosome density (fig. S10C). All these features promote R-loop formation.

The association of AtNDX with the *COOLAIR* promoter–*FLC* terminator (Fig. 2A) and its ssDNA binding capacity led us to test whether R-loops form in this region. We used native sodium bisulfite treatment, which can convert cytosine (C) to uracil (U) if the C bases are located on the unprotected ssDNA strand (Fig. 3A and fig. S11A) (*18*). The mutation profile of the nontemplate ssDNA region allows definition of the position and length of the R-loop. We found C to U conversion only on the noncoding DNA strand [Fig. 3C and figs. S11 and S12; thymine (T) was detected after polymerase chain reaction (PCR) amplification], indicating that R-loops are formed by antisense transcripts. The total length of the R-loop region is about 300 to 700 bp, with a well-defined 5′ end starting 200 bp upstream of the multiple *COOLAIR* transcription start sites (*3*, *20*). This overlaps with the *FLC* region that we have previously defined as a heterochromatic patch containing Histone H3 dimethyl Lys^9 (H3K9me2) and homologous small RNAs (*20*). The 3′ end of the R-loop is more variable, terminating 100 to 500 bp downstream of the *COOLAIR* transcription start site window (Fig. 3C) with the longer forms terminating in the region of the *COOLAIR* proximal polyadenylation site (Fig. 3C). The heterogeneity at the 3′ limit of the R-loop may be caused by different rates of pausing and polymerase drop-off during transcription elongation (*16*, *21*), or more likely cotranscriptional RNA processing (*4*) with factors such as SR proteins resolving the RNA-DNA hybrids (*22*).

We speculated that AtNDX could play a role in R-loop formation or stabilization on the basis of its ssDNA binding capacity (Fig. 2B) and its chromatin association (Fig. 2A). We therefore compared R-loop formation in *Col*-0 and *eoc1* by DNA immunoprecipitation (DIP) using the RNA-DNA hybrid–specific antibody S9.6 (*15*, *17*, *23*). We found that the DIP signal was enriched over the *COOLAIR* promoter (Fig. 3D), was sensitive to ribonuclease H (fig. S13), and was reduced by a factor of ~3 in *eoc1* (Fig. 3D, region i). This suggests that the R-loop naturally forms over the *COOLAIR* promoter, after which AtNDX binds to the displaced nontemplate ssDNA, thereby stabilizing the R-loop structure (Fig. 4D). AtNDX binding may hamper the accessibility of RNA-DNA helicases required to resolve the R-loop, which in turn could affect Pol II initiation and/or elongation (*21*). The presence of the R-loop and AtNDX binding would then affect initiation and/or elongation of *COOLAIR* transcription. Nuclear run-on data confirmed that stabilization of the R-loop reduced transcription of the endogenous *COOLAIR* as well as the reporter *Luc* fusion (Fig. 4A and fig. S14A). R-loops formed at transcriptional termination sites can affect Pol II pausing and read-through (*15*, *24*). We therefore tested whether the R-loop and AtNDX binding also affected transcription termination of *FLC*. *FLC* read-through transcription is unchanged in *eoc-1* relative to the wild type (Fig. 4B, regions j, k, and m). However, a role for R-loops in *FLC* sense transcription termination cannot be excluded, because the R-loop is not fully disrupted in *eoc1* (Fig. 3D).

We next addressed how R-loop stabilization and repression of *COOLAIR* might influence *FLC* gene expression. *eoc1* did not increase *FLC* expression when combined with autonomous pathway mutants, which suggests that reduced stabilization of the R-loop increases *FLC* expression via the involvement of *COOLAIR* in the autonomous pathway mechanism (Fig. 4C). We also asked whether AtNDX changed *FLC* regulation via perturbation of the gene loop generated through physical interaction of the 5′ and 3′ flanking regions of *FLC* (*25*). A robust gene loop was still detected in *eoc1*, suggesting independent activities (fig. S15). Previous analysis had identified a small patch of heterochromatin marked by H3K9me2 and homologous small RNAs immediately downstream of the *FLC* sense transcript polyadenylation site (*20*). The small RNAs, which are dependent on the alternative plant RNA polymerase Pol IV (*20*), were still detected in *eoc1* mutants (fig. S16).

R-loops were initially thought to be a rare byproduct of transcription but more recently have been found to cause genome instability (*13*). Our findings indicate that AtNDX homeodomain R-loop stabilization is an important factor regulating expression of the lncRNA *COOLAIR*.

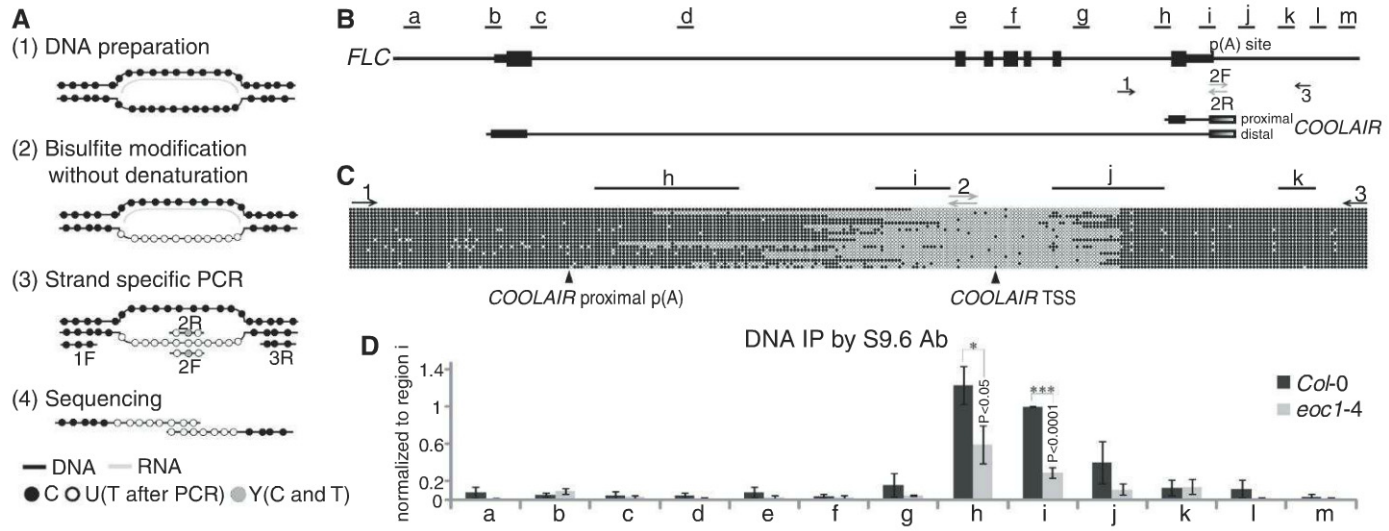


Fig. 3. AtNDX stabilizes the R-loop in the endogenous *FLC* terminator. (**A**) Schematic steps of the R-loop footprinting method. C's (black dots) in ssDNA can be converted to U's (white dots) by bisulfite only in nonprotected DNA (fig. S11A). The length of the R-loop is revealed after strand-specific PCR and sequencing. (**B**) Illustration of *FLC* gene structure. Amplicons from regions a to m were used in qPCR for ChIP and DIP. Primers 1 and 3 were native primers (black); primers 2F and 2R (gray) carried C to T conversions. (**C**) Sequence analysis of individual clones showing R-loop footprinting amplified by 1/2R and 2F/3 primer pairs. Each circle indicates a C nucleotide; empty circles indicate T's converted from C's, and solid circles indicate unconverted C's. (**D**) DNA immunoprecipitation using RNA-DNA hybrid–specific antibody S9.6. Values of DNA enrichment were divided by input and normalized to region i shown in (B); error bars are SEM of four biological replicates. The amplicons corresponding to R-loop region are shown at the top of (C). $^*P < 0.05$, $^{***}P < 0.001$.

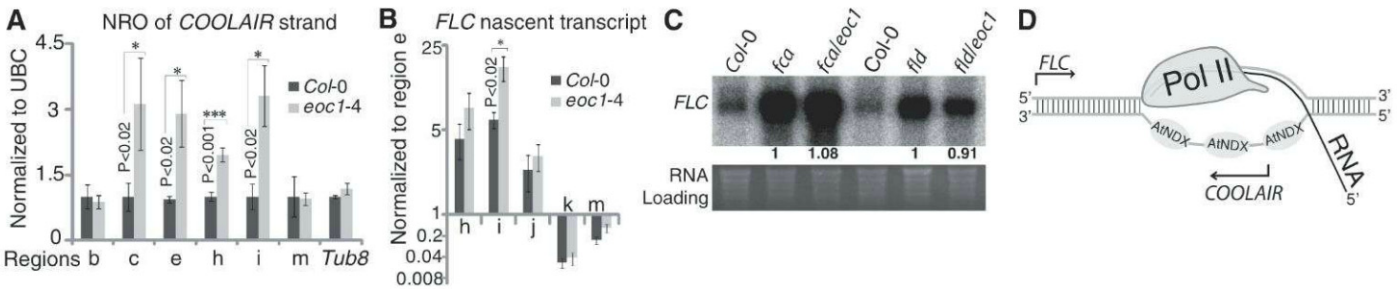


Fig. 4. Reduced R-loop formation releases transcriptional repression of *COOLAIR*. (**A**) Strand-specific nuclear run-on (NRO) analysis (*26*) of transcription rate of *COOLAIR*. Error bars are SEM from four biological replicates. (**B**) Quantification of nascent and read-through transcripts of sense *FLC* by quantitative reverse transcription PCR. Error bars are SEM of nine biological replicates. Data were normalized to *UBC21* and then to region e. (**C**) *FLC* expression in the histone K4 demethylase mutant *fld* (*7*), *fld/eoc1*, the RNA-binding protein mutant *fca* (*4*), and *fca/eoc1* assessed by Northern blotting. (**D**) Model of how the R-loop influences lncRNA *COOLAIR* expression. AtNDX stabilizes the R-loop by binding to an ssDNA strand. This may lead to Pol II stalling and/or abortion of transcription (*21*), or it could affect cotranscriptional RNA metabolism (*4*, *22*). $^*P < 0.05$, $^{***}P < 0.001$.

COOLAIR regulation is also influenced by other pathways quantitatively regulating *FLC* expression and flowering (*8*). Stabilization of the R-loop thus provides an additional regulatory layer contributing to the robustness of flowering time regulation. R-loop stabilization by ssDNA binding proteins may be a general mechanism influencing gene expression in many organisms.

References and Notes

1. C. Shindo, C. Lister, P. Crevillen, M. Nordborg, C. Dean, *Genes Dev.* **20**, 3079 (2006).
2. V. Coustham *et al.*, *Science* **337**, 584 (2012).
3. S. Swiezewski, F. Liu, A. Magusin, C. Dean, *Nature* **462**, 799 (2009).
4. F. Liu, S. Marquardt, C. Lister, S. Swiezewski, C. Dean, *Science* **327**, 94 (2010).
5. A. Angel, J. Song, C. Dean, M. Howard, *Nature* **476**, 105 (2011).
6. J. Song, A. Angel, M. Howard, C. Dean, *J. Cell Sci.* **125**, 3723 (2012).
7. F. Liu *et al.*, *Mol. Cell* **28**, 398 (2007).
8. R. Ietswaart, Z. Wu, C. Dean, *Trends Genet.* **28**, 445 (2012).
9. J. L. Rinn, H. Y. Chang, *Annu. Rev. Biochem.* **81**, 145 (2012).
10. K. Mukherjee, L. Brocchieri, T. R. Bürglin, *Mol. Biol. Evol.* **26**, 2775 (2009).
11. J. E. Jørgensen *et al.*, *Plant Mol. Biol.* **40**, 65 (1999).
12. R. Rivera-Pomar, D. Niessing, U. Schmidt-Ott, W. J. Gehring, H. Jäckle, *Nature* **379**, 746 (1996).
13. A. Aguilera, T. García-Muse, *Mol. Cell* **46**, 115 (2012).
14. A. El Hage, S. L. French, A. L. Beyer, D. Tollervey, *Genes Dev.* **24**, 1546 (2010).
15. K. Skourti-Stathaki, N. J. Proudfoot, N. Gromak, *Mol. Cell* **42**, 794 (2011).
16. P. Huertas, A. Aguilera, *Mol. Cell* **12**, 711 (2003).
17. P. A. Ginno, P. L. Lott, H. C. Christensen, I. Korf, F. Chédin, *Mol. Cell* **45**, 814 (2012).
18. K. Yu, F. Chedin, C. L. Hsieh, T. E. Wilson, M. R. Lieber, *Nat. Immunol.* **4**, 442 (2003).
19. E. J. Finnegan *et al.*, *Plant J.* **44**, 420 (2005).
20. S. Swiezewski *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 3633 (2007).
21. A. G. Rondón, S. Jimeno, A. Aguilera, *Biochim. Biophys. Acta* **1799**, 533 (2010).
22. X. Li, J. L. Manley, *Cell* **122**, 365 (2005).
23. S. J. Boguslawski *et al.*, *J. Immunol. Methods* **89**, 123 (1986).
24. N. J. Proudfoot, *Trends Biochem. Sci.* **14**, 105 (1989).
25. P. Crevillén, C. Sonmez, Z. Wu, C. Dean, *EMBO J.* **32**, 140 (2013).
26. F. Liu, S. Bakht, C. Dean, *Science* **335**, 1621 (2012).

Acknowledgments: We thank all the members of the Dean lab for useful discussions; G. Calder and S. Rosa for help with microscopy; and the Nottingham Arabidopsis Stock Centre (NASC) and Institut National de la Recherche Agronomique (INRA) for *Arabidopsis* lines. Supported by a Wellcome Trust programme grant (K.S.-S. and N.J.P.). The Dean lab is supported by the UK Biotechnology and Biological Sciences Research Council (BBSRC) and a European Research Council advanced investigator grant. C.D. holds stock in Mendel Biotechnology.

Supplementary Materials

www.sciencemag.org/cgi/content/full/340/6132/619/DC1
Materials and Methods
Figs. S1 to S16
Tables S1 to S4
References (*27*–*32*)

7 January 2013; accepted 12 March 2013
10.1126/science.1234848

Targeted Inhibition of Mutant IDH2 in Leukemia Cells Induces Cellular Differentiation

Fang Wang,[1*] Jeremy Travins,[1*] Byron DeLaBarre,[1*] Virginie Penard-Lacronique,[2,3,4*] Stefanie Schalm,[1*] Erica Hansen,[1] Kimberly Straley,[1] Andrew Kernytsky,[1] Wei Liu,[1] Camelia Gliser,[1] Hua Yang,[1] Stefan Gross,[1] Erin Artin,[1] Veronique Saada,[3] Elena Mylonas,[2,3,4] Cyril Quivoron,[2,3,4] Janeta Popovici-Muller,[1] Jeffrey O. Saunders,[1†] Francesco G. Salituro,[1‡] Shunqi Yan,[5] Stuart Murray,[1] Wentao Wei,[6] Yi Gao,[7] Lenny Dang,[1] Marion Dorsch,[1] Sam Agresta,[1] David P. Schenkein,[1] Scott A. Biller,[1] Shinsan M. Su,[1] Stephane de Botton,[2,3,4] Katharine E. Yen[1§]

A number of human cancers harbor somatic point mutations in the genes encoding isocitrate dehydrogenases 1 and 2 (IDH1 and IDH2). These mutations alter residues in the enzyme active sites and confer a gain-of-function in cancer cells, resulting in the accumulation and secretion of the oncometabolite (*R*)-2-hydroxyglutarate (2HG). We developed a small molecule, AGI-6780, that potently and selectively inhibits the tumor-associated mutant IDH2/R140Q. A crystal structure of AGI-6780 complexed with IDH2/R140Q revealed that the inhibitor binds in an allosteric manner at the dimer interface. The results of steady-state enzymology analysis were consistent with allostery and slow-tight binding by AGI-6780. Treatment with AGI-6780 induced differentiation of TF-1 erythroleukemia and primary human acute myelogenous leukemia cells in vitro. These data provide proof-of-concept that inhibitors targeting mutant IDH2/R140Q could have potential applications as a differentiation therapy for cancer.

Somatic point mutations affecting one of three active-site arginine residues of isocitrate dehydrogenase (IDH1/R132, IDH2/R140, and IDH2/R172) define distinct subsets of low-grade glioma and secondary glioblastoma, chondrosarcoma, cholangiocarcinomas, and acute myelogenous leukemia (AML) (*1*). IDH is a metabolic enzyme that interconverts isocitrate and α-ketoglutarate (αKG), but cancer-associated point mutations in IDH1 and IDH2 confer a neomorphic activity that allows reduction of αKG to the oncometabolite (*R*)-2-hydroxyglutarate (2HG) (*2*). High concentrations of 2HG have been shown to inhibit αKG-dependent dioxygenases, including histone and DNA demethylases, which play a key role in regulating the epigenetic state of cells (*3*–*5*). Patients harboring IDH mutations display a CpG island methylator phenotype (CIMP), and overexpression of IDH mutant enzymes can induce histone and DNA hypermethylation, as well as block cellular differentiation (*6*–*8*). In addition, mice engineered to express IDH1/R132H in hematopoietic tissue display an increased number of early hematopoietic progenitors, as well as splenomegaly, anemia, hypermethylated histones, and altered DNA methylation patterns similar to those found in AML patients harboring IDH1/2 mutations (*9*). Together, these data suggest that cancer-associated IDH mutations may induce a block in cellular differentiation to promote tumorigenesis.

To elucidate the relationship between mutant enzyme activity, 2HG levels, and oncogenic state, we developed a small molecule that selectively inhibits the IDH2/R140Q mutant expressed in

[1]Agios Pharmaceuticals, Cambridge, MA, 02139–4169, USA. [2]Institut National de la Santé et de la Recherche Médicale, INSERM U985, F-94805 Villejuif, France. [3]Institut Gustave Roussy, F-94805 Villejuif, France. [4]Université Paris-Sud, F-91405 Orsay, France. [5]Schrödinger, Inc., 120 West 45th Street, New York, NY 10036, USA. [6]Viva Biotech, 334 Aidisheng Rd, Zhangjiang High-tech Park, Shanghai, 201201, China. [7]Shanghai ChemPartner Co., 998 Halei Road, Pudong New Area, Shanghai, 201203, China.

*These authors contributed equally to this work.
†Present address: Ember Therapeutics, 480 Arsenal Street, Watertown, MA 02347, USA.
‡Present address: Sage Therapeutics, 215 First Street, Cambridge, MA 02141, USA.
§Corresponding author. E-mail: katharine.yen@agios.com

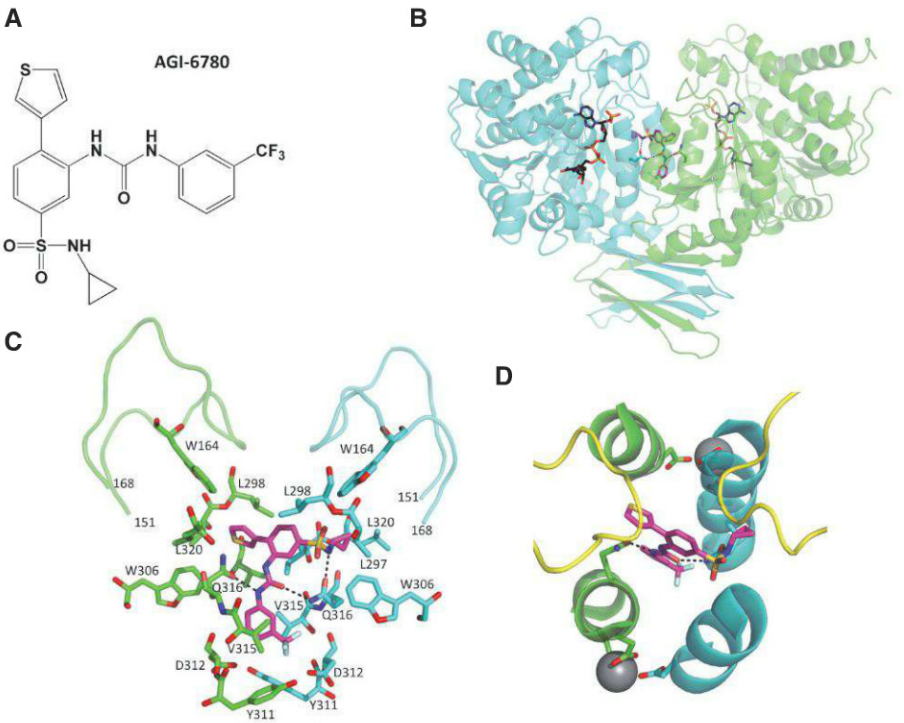


Fig. 1. Chemical structure of AGI-6780 (A) and co-complex of AGI-6780 and NADP$^+$ with IDH2/R140Q (B). Individual chains are colored in light blue and green. A single molecule of AGI-6780 binds at the dimer interface. NADP$^+$ molecules are shown with carbons in black, AGI-6780 is shown with carbons in magenta, and Gln^{316} is shown with carbon chains matching their respective A and B chains. (**C**) Detailed view of the environment around AGI-6780. The loops (residues 151 to 168) that fold over the open end of the binding pocket are depicted in ribbon form. Hydrogen bonds are depicted as black dashes. (**D**) A top-down view of the binding pocket showing the 151–168 loops folded over the top. The green helices and the loop immediately above are from one monomer and the blue helices with the loop immediately above are from the opposite monomer. The 151–168 loops are shown in yellow and fold over the open end of the binding pocket. The Gln^{316} residues that hydrogen bond to AGI-6780 are shown in stick form with H-bonds depicted as dashed lines. The kinetically inert calcium ions that occupy the site of the catalytic divalent metal ions are shown as gray spheres. The metal ion is ligated by Asp^{314} and Asp^{318} from one monomer and Asp^{291} from the opposite monomer, with the rest of the coordination sphere being completed by water molecules when in the open conformation. The equivalent Asp^{314}/Asp^{318}-containing helix in cytosolic IDH1 was observed to unfold in the open conformation. Abbreviations for the amino acid residues are as follows: D, Asp; Lys; L, Leu; Q, Gln; V, Val; W, Trp; and Y, Tyr.

the leukemia cells of ~9% of AML patients (*10*). Screening of a small-molecule library using an enzyme assay in the presence of saturating nicotinamide adenine dinucleotide phosphate (NADPH) identified a series of heterocyclic urea sulfonamides as inhibitors of IDH2/R140Q. Early x-ray crystallographic studies of exemplary active small molecules complexed with IDH2/R140Q provided evidence of a single inhibitor molecule binding at an allosteric site located well within the dimer interface. Kinetic characterization revealed these urea sulfonamides to be slow-tight binders that exhibit noncompetitive inhibition with respect to substrate (fig. S1A) and uncompetitive inhibition with respect to the NADPH cofactor (fig. S1B). Medicinal chemistry optimization of the urea sulfonamide series led to the design of AGI-6780 (Fig.1A), a molecule with nanomolar potency for 2HG inhibition and long residence time (k_{on} = 5.8×10^4 M^{-1} min^{-1}; k_{off} = 8.3×10^{-3} min^{-1}) (fig. S1, C and D). This time-dependent inhibition activity of AGI-6780 favors the mutated enzyme because it was less potent against IDH2-WT (wild type) (Table 1). Furthermore, AGI-6780 displayed excellent potency in reducing 2HG levels in cell lines ectopically overexpressing IDH2/R140Q [median effective concentration (EC_{50}) = <20 nM], as well as excellent selectivity against other dehydrogenases, including the closely related IDH1-WT or R132H mutant enzymes (Table 1). The tight-binding, allosteric mode of inhibition of AGI-6780 is consistent with the observation that similar drug concentrations can inhibit mutant IDH2 comparably in enzymatic- and cellular-based 2HG assays.

To confirm that the AGI-6780 binding mode was similar to that observed with the initial active molecules, we determined a high-resolution crystal structure of homodimeric IDH2/R140Q complexed with $NADP^+$, Ca^{2+}, and AGI-6780 (Fig. 1B and table S1). As expected, a single inhibitor molecule was observed at the central IDH2 dimer interface created by four helices, residues 290 to 299 and 310 to 322 from each of the A and B chains. The AGI-6780 binding site is opposite to the side of the helices that project aspartate residues necessary for the chelation of a catalytic divalent ion such as Mg^{2+}. The highly symmetric AGI-6780 binding pocket extends deep into the protein interface and is closed over by loops composed of residues 152 to 167 (Fig. 1, C and D). These loops have a slightly higher temperature factor than the surrounding residues (Fig. 1D), suggesting increased mobility. They also fold over the pocket and would need to move to allow access into or out of the allosteric site, providing an explanation for the slow-tight binding of AGI-6780 to IDH2/R140Q described above. AGI-6780 makes several direct hydrogen bonding interactions from its urea group and amide nitrogen to Gln^{316} (Fig. 1C and fig. S2B), but a large amount of binding arises from van der Waals contacts between the binding site on the protein and the more hydrophobic regions of AGI-6780. An open conformation for the protein was observed whereby the large cofactor binding domain (residues 41 to 158 and 324 to 456) is separated from the small catalytic domain (residues 159 to 181 and 224 to 323) and is nearly identical (<1 Å root mean square deviation) to that observed for the yeast mitochondrial IDH determined in the absence of substrate molecules (fig. S2A) (*11*). This open conformation is analogous to that observed for cytosolic IDH1 (*12*) with the major exception that the catalytic metal-coordinating helices remain intact in IDH2 (Fig. 1D) whereas they are unfolded in the IDH1 open conformation. Because IDH1 and IDH2 have different allosteric sites, highly selective targeting of the mitochondrial IDH2 isoform was achieved with AGI-6780.

Our approach to designing inhibitors against IDH2 with slow-tight binding and noncompetitive inhibitory properties arose from the hypothesis that antitumor efficacy would require 2HG to be fully suppressed to background levels under conditions where cellular cofactor concentrations may far exceed IDH2 mutant intrinsic affinity for NADPH ($K_{M,\ NADPH}$ = 200 nM). The discovery of heterocyclic urea sulfonamides that bind in a unique allosteric pocket of IDH2/R140Q confirmed this hypothesis. The binding of AGI-6780 at the dimer interface suggests that its role as an allosteric inhibitor arises from its ability to hold the IDH2/R140Q dimer in an open conformation that is incompatible with catalysis. The allosteric site where AGI-6780 is bound does not directly involve Gln^{140}. Nonetheless, AGI-6780 exhibits excellent potency against the neomorphic activity of both the mutant homodimer and IDH2/WT:IDH2/R140Q heterodimeric enzymes, although exhibiting less potency against the normal oxidative decarboxylation activity of IDH2-WT (Table 1).

The inhibitory activity against both mutant and wild-type subunits could offer an advantage in the context of heterozygous IDH mutations because studies of the IDH1/WT: IDH1/R132H heterodimer suggested that the wild-type and mutant subunits contribute to the production of 2HG either in concert or independently, depending on whether the carbon source is derived from isocitrate or from αKG, respectively (*13*). Genetic knockdown of IDH1/WT alone decreased 2HG production (*14*), which further underscores the contribution of the wild-type IDH allele to 2HG production in some cancer cells. Overall, we believe that both the allosteric nature observed by enzyme kinetics and crystallography and the slow-tight binding characteristics measured with enzyme kinetics are essential features of AGI-6780 for optimal suppression of 2HG production arising from the IDH2 mutation in cancer cells.

To study the biological effect of AGI-6780, we clonally selected mutant IDH2/R140Q protein

Table 1. Biochemical and cellular potency for 2HG inhibition by AGI-6780. The time-dependent inhibition against homodimeric or heterodimeric IDH2/WT or IDH2/R140Q mutant by AGI-6780 was investigated after either 1 or 16 hours of preincubation. AGI-6780 was tested in both human glioblastoma U87 and TF-1 cells expressing IDH2/R140Q, as well as against IDH1/R132H. Selectivity against other dehydrogenases is also noted. Selectivity against lactate dehydrogenase (LDHA), 3-phosphoglycerate dehydrogenase (3PGDH), glutamate dehydrogenase (GDH), and glucose-6-phosphate dehydrogenase (G6PDH) is also noted; values are the mean of at least three determinations ± SD.

AGI-6780 biochemical properties, IC_{50} for α-KG reduction		
Enzymes assayed	*Incubation time (hours)*	*IC_{50} (nM)*
IDH2-R140Q	1	170 ± 47
IDH2-R140Q	16	23 ± 1.7
IDH2-R140Q/WT	1	120 ± 42
IDH2-R140Q/WT	16	4.0 ± 1.2
IDH1-R132H	16	11000 ± 84
AGI-6780 biochemical properties, IC_{50} for isocitrate to α-KG		
Enzymes assayed	*Incubation time (hours)*	*IC_{50} (nM)*
IDH2-WT	1	2700 ± 31
IDH2-WT	16	190 ± 8.1
IDH1-WT	16	>100000
AGI-6780 cellular properties, IC_{50} for 2-HG formation		
Cell line	*Incubation time (hours)*	*IC_{50} (nM)*
TF-1 (IDH2-R140Q)	48	18 ± 0.51
U87 (IDH2-R140Q)	48	11 ± 2.6
U87 (IDH1-R132H)	48	>100000
AGI-6780 dehydrogenase selectivity panel		
Enzymes assayed	*Incubation time (hours)*	*IC_{50} (nM)*
LDHA (pyruvate to lactate)	1	>100000
LDHA (lactate to pyruvate)	1	>100000
3PGDH	1	>100000
GDH	1	>100000
G6PDH	1	>100000

in granulocyte-macrophage colony-stimulating factor (GM-CSF)–dependent human TF-1 erythroleukemia cells (*15*). These cells (TF1/R140Q) produce large amounts of 2HG similar to that observed in human tumors (5 to 30 mM), which correlates with mutant IDH2 expression and growth in the absence of GM-CSF (Fig. 2A and fig. S3) (*16*, *17*). Expression of IDH2/R140Q also induced a morphological change in TF-1 cells: They became more spindle in shape and attached to the tissue culture plate, two features often found in undifferentiated mesenchymal or "stem" cells (Fig. 2B) (*18*). Similarly, immunoblot analysis showed an increase in vimentin expression (Fig. 2C), a marker found in human hematopoietic stem cells (HSCs), but not in cells differentiating along the erythrocyte lineage (*19*). Together, these data suggest that IDH2/R140Q expression can induce a more immature phenotype. To test this possibility, we analyzed cell surface expression of CD34 and CD38 by flow cytometry. CD34 expression is found on human hematopoietic stem and progenitor cells, and high CD38 expression is only seen at later stages of hematopoietic differentiation. Although CD34 and CD38 are both found on the cell surface of TF-1/pLVX vector cells, CD38 expression was decreased in the IDH2/R140Q cells, suggesting a more primitive state of differentiation (Fig. 2D). These data demonstrate that IDH2/R140Q can either shift TF-1 cells toward an earlier stage and/or block hematopoietic cell differentiation and is similar to the results found with the IDH1/R132H mutation (*20*).

Because IDH mutations are reported to induce a block in hematopoietic cell differentiation (*6*, *9*), we tested whether IDH2/R140Q expression could block erythropoietin (EPO)–induced differentiation of TF-1 cells. Unlike TF-1/pLVX cells, EPO treatment of TF-1/R140Q cells failed to induce differentiation, as evidenced by the lack of color change associated with hemoglobin expression (Fig. 2E). Next, we looked at the expression of hemoglobin gamma (*HBG*) and Kruppel-like factor 1 (*KLF1*), as both genes are known to be upregulated during erythropoiesis (*21*). We found that EPO induced the expression of both genes in vector cells but not in IDH2/R140Q cells with high concentrations of 2HG (Fig. 2, F and G). However, treatment of TF-1/R140Q cells with AGI-6780, at concentrations that lowered 2HG to near-normal physiologic levels, restored expression of both *HBG* and *KLF1* genes and the color change associated with differentiation (Fig. 2, E to G).

To investigate the effect of AGI-6780 on primary human AML cells ex vivo, we treated patient samples containing either the IDH2/R140Q mutant ($n = 4$) or IDH2/WT ($n = 5$) (table S2). Patient's blood or bone marrow samples were sorted (Sytox staining) and cultured in condi-

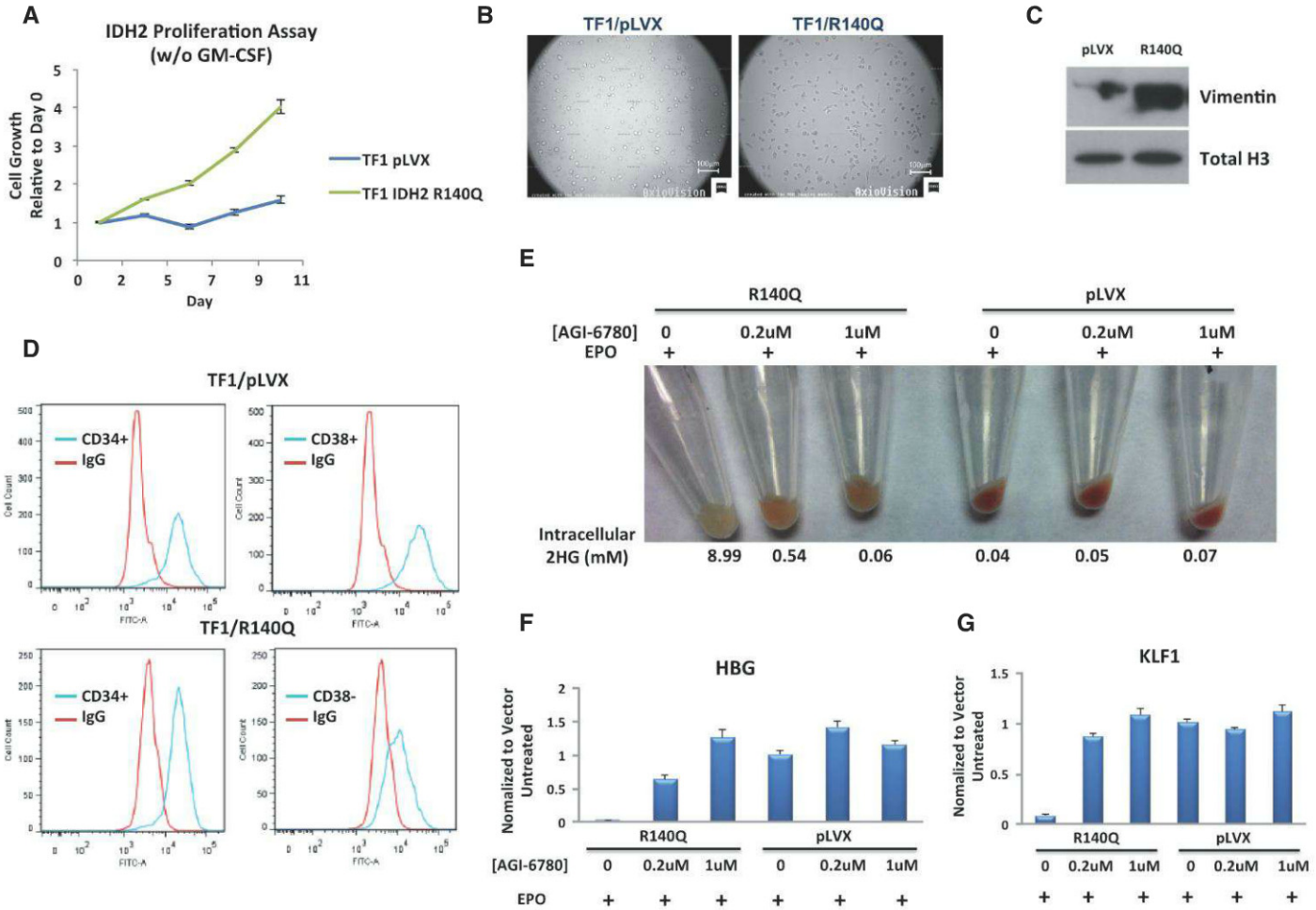


Fig. 2. AGI-6780 can reverse the IDH2/R140Q-induced differentiation block in TF-1 cells. (**A**) Expression of IDH2/R140Q in TF-1 cells confers GM-CSF–independent growth. (**B**) IDH2/R140Q expression induces morphological changes in TF-1 cells where the cells become more spindle in shape and attach to the tissue culture plate. (**C**) Vimentin expression is significantly induced by IDH2/R140Q expression in TF-1 cells. (**D**) Cells expressing IDH2/R140Q have decreased CD38 cell surface expression as measured by flow cytometry analysis; no effect on CD34 expression was observed. (**E**) Treatment of TF1/pLVX cells with EPO induces erythroid differentiation (color change induced by hemoglobin expression), whereas TF1/R140Q cells were resistant to EPO-induced differentiation. Pretreatment with AGI-6780 markedly decreased the intracellular concentration of 2HG in the TF1/R140Q cells and restored their ability to undergo EPO-induced differentiation. Pretreatment of AGI-6780 restored the mRNA expression of (**F**) hemoglobin G1/2 and (**G**) KLF1 in the TF1/R140Q cells in the presence of EPO. Hemoglobin G1/2 and KLF1 expression were analyzed by real-time polymerase chain reaction and values were normalized to dimethyl sulfoxide (DMSO)–treated pLVX control.

tioned medium either in the presence or absence of a single dose of AGI-6780, and samples from each culture were collected over a 2-week time period (days 1, 3, 8, 10, and 13). IDH2 mutant samples treated with AGI-6780 showed a dose-dependent reduction in the amounts of extracellular (Fig. 3A and fig. S4A) and intracellular (Fig. 3B) 2HG. In the presence AGI-6780, a burst of cell proliferation was also observed only in the IDH2/R140Q mutant samples starting at day 4 of culture (Fig. 3C). Furthermore, AGI-6780 did not exert a cytotoxic effect as the number of viable, more mature CD45-positive cells increased in the presence of AGI-6780 only in the IDH2/R140Q mutant samples whereas the IDH2/WT samples remained unchanged (fig. S4B). In these studies, blasts were identified by flow cytometry analysis as $CD45^{low}$ cells with low side scatter (SSC^{low}) using a SSC versus CD45 plot. After 6 days of inhibitor treatment, a dose-dependent decrease in the percentage of $CD45^{low}$ blast cells was observed in the IDH2/R140Q samples but not in IDH2/WT blasts (fig. S5A).

In parallel, maturation of AML blasts was also evaluated by flow cytometry analysis for changes in cell surface markers associated with monocytic and granulocytic differentiation (CD14, CD15, and CD11b) along with intracellular myeloperoxidase (MPO) as an indicator of maturation into the neutrophilic pathway. Although all samples (wild-type and IDH2/R140Q) underwent some level of spontaneous differentiation as a result of the cell-culturing conditions (i.e., cytokine addition), treatment with AGI-6780 induced an increase in the number of differentiated cells only in IDH2/R140Q patient samples. In the presence of AGI-6780, an increase in the mean fluorescence intensity for all cell surface markers and MPO could be seen in the IDH2 mutant samples, but not in wild-type IDH2 samples (Fig. 3D). Similarly, cytology revealed that by day 8 to 9 of treatment, the percentage of blasts or myeloblasts had decreased, and early signs of maturation were observed by the dose-dependent increase in the number of myelocytes and metamyelocytes (fig. S5, B to D). These effects on differentiation were only seen in mutant samples.

To establish that patient samples carrying the IDH2/R140Q allele were not counterselected during the in vitro differentiation assays, we cultured unfractionated nucleated bone marrow cells from patient #3 (IDH2/R140Q) for 13 days in methylcellulose-containing growth factors (stem cell factor, GM-CSF, interleukin-3, and erythropoietin) in the presence or absence of 5 μM AGI-6780. In the presence of AGI-6780, cells showed a threefold increase in the number of colony-forming units when compared to the number formed in vehicle-treated cells (fig. S5E). We determined the proportion of the IDH2/R140Q mutant among the colonies to eliminate the participation of residual normal progenitors in this

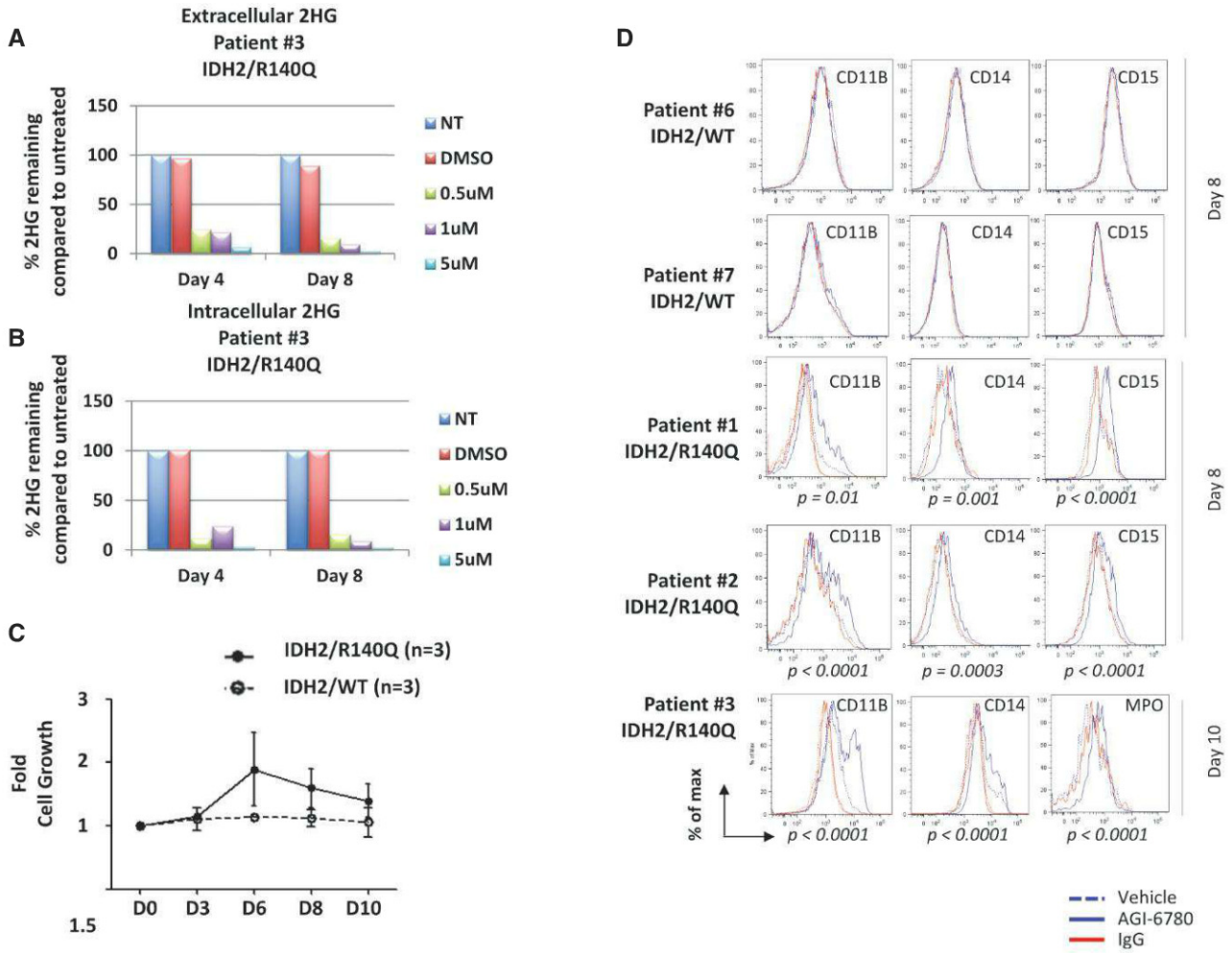


Fig. 3. AGI-6780 can induce blast differentiation in primary human IDH2/R140Q AML patient samples. A dose-dependent decrease in extracellular (**A**) and intracellular (**B**) 2HG was observed from IDH2/R140Q patient samples treated with AGI-6780. NT: no treatment. (**C**) The absolute number of cells increases upon treatment with AGI-6780 (5 μM) in IDH2/R140Q patient samples but not in IDH2/WT patient samples. Values were normalized to vehicle-treated cells and data are presented as mean ± SEM IDH2/R140Q (n = 3; closed circles, patients #1 to 3) and IDH2/WT (n = 3; open circles, patients #5 to 7) samples. (**D**) Flow cytometry analysis shows an increase in CD11b-, CD14-, CD15-, and MPO-positive cells in IDH2/R140Q patient samples but not in IDH2/WT samples after treatment with AGI-6780. The association between AGI-6780 treatment and cell surface maturation markers expression was tested with the Chi-square test with Yates' continuity correction; P values are indicated.

process. The heterozygous IDH2/R140Q mutation was present in all of the colonies cultured with or without AGI-6780, confirming that the inhibitor promotes outgrowth and differentiation of cells harboring the mutated allele (fig. S5F). Similar results were obtained with patient #3's circulating leukemic cells cultured in methylcellulose and in liquid media. Taken together, these data demonstrate that targeted inhibition of IDH2/R140Q by AGI-6780 can specifically induce differentiation of AML blasts derived from patients with the IDH2/R140Q mutation.

In this study, we have used GM-CSF–dependent erythroleukemia TF-1 cells expressing IDH2/R140Q and a tool compound, AGI-6780, as a model system for evaluating the potential utility of IDH2 mutant inhibitors in the treatment of leukemias expressing the IDH2/R140Q mutant. Expression of mutant IDH2 in these cells conferred growth factor–independent proliferation and induced a block in erythroid differentiation that was reversed with AGI-6780 treatment. Furthermore, ex vivo treatment of primary human IDH2/R140Q mutant AML cells with AGI-6780 resulted in differentiation of the AML blasts down the macrophage and granulocytic lineages. These data are reminiscent of the leukemic cell differentiation response observed upon exposure to *all-trans* retinoic acid (ATRA) in acute promyelocytic leukemia (*22*) and demonstrate that inhibition of mutant IDH2 can relieve a block in differentiation present in this leukemic subset. Finally, these data support the clinical evaluation of IDH2 mutant–targeted agents in AML and other malignancies.

References and Notes

1. K. E. Yen, M. A. Bittinger, S. M. Su, V. R. Fantin, *Oncogene* **29**, 6409 (2010).
2. L. Dang *et al.*, *Nature* **462**, 739 (2009).
3. R. Chowdhury *et al.*, *EMBO Rep.* **12**, 463 (2011).
4. P. Koivunen *et al.*, *Nature* **483**, 484 (2012).
5. W. Xu *et al.*, *Cancer Cell* **19**, 17 (2011).
6. C. Lu *et al.*, *Nature* **483**, 474 (2012).
7. S. Turcan *et al.*, *Nature* **483**, 479 (2012).
8. M. E. Figueroa *et al.*, *Cancer Cell* **18**, 553 (2010).
9. M. Sasaki *et al.*, *Nature* **488**, 656 (2012).
10. P. S. Ward *et al.*, *Cancer Cell* **17**, 225 (2010).
11. Y. Peng, C. Zhong, W. Huang, J. Ding, *Protein Sci.* **17**, 1542 (2008).
12. B. Yang, C. Zhong, Y. Peng, Z. Lai, J. Ding, *Cell Res.* **20**, 1188 (2010).
13. B. Pietrak *et al.*, *Biochemistry* **50**, 4804 (2011).
14. G. Jin *et al.*, *Cancer Res.* **73**, 496 (2013).
15. E. Pronier *et al.*, *Blood* **118**, 2551 (2011).
16. S. Gross *et al.*, *J. Exp. Med.* **207**, 339 (2010).
17. S. H. Fathi *et al.*, *Blood* **120**, 4649 (2012).
18. R. Huss, C. Lange, E. M. Weissinger, H. J. Kolb, K. Thalmeier, *Stem Cells* **18**, 252 (2000).
19. R. Majeti, C. Y. Park, I. L. Weissman, *Cell Stem Cell* **1**, 635 (2007).
20. J.-A. Losman *et al.*, *Science* **339**, 1621 (2013).
21. M. J. Koury, S. T. Sawyer, S. J. Brandt, *Curr. Opin. Hematol.* **9**, 93 (2002).
22. Z. Y. Wang, Z. Chen, *Blood* **111**, 2505 (2008).

Acknowledgments: We thank P. Rameau for flow and cell-sorting cytometry; E. Leclercq, M.-H. Courtier, and A. Bennaceur for NPM1 and FLT3 mutational status analyses; W. G. Kaelin and J. Aurore-Losman for sharing unpublished data on the biology of TF-1 cells and IDH mutations; and I. Mellinghoff, members of the Mellinghoff laboratory, M. Vander Heiden, L. Cantley, and J. Engelman for thoughtful discussions over the course of this work. Agios Pharmaceuticals has filed a patent application on AGI-6780. Atomic coordinates and structure factors have been deposited in the Protein Data Bank with identification code 4JA8.

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1234769/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 and S2
References (*23*–*33*)

4 January 2013; accepted 5 March 2013
Published online 4 April 2013;
10.1126/science.1234769

An Inhibitor of Mutant IDH1 Delays Growth and Promotes Differentiation of Glioma Cells

Dan Rohle,[1,2]* Janeta Popovici-Muller,[3]* Nicolaos Palaskas,[1]* Sevin Turcan,[1]* Christian Grommes,[4] Carl Campos,[1] Jennifer Tsoi,[8] Owen Clark,[1] Barbara Oldrini,[1] Evangelia Komisopoulou,[8] Kaiko Kunii,[3] Alicia Pedraza,[7] Stefanie Schalm,[3] Lee Silverman,[3] Alexandra Miller,[4] Fang Wang,[3] Hua Yang,[3] Yue Chen,[3] Andrew Kernytsky,[3] Marc K. Rosenblum,[6] Wei Liu,[3] Scott A. Biller,[3] Shinsan M. Su,[3] Cameron W. Brennan,[1,7] Timothy A. Chan,[1,5] Thomas G. Graeber,[8]† Katharine E. Yen,[3]†‡ Ingo K. Mellinghoff[1,2,4]†‡

The recent discovery of mutations in metabolic enzymes has rekindled interest in harnessing the altered metabolism of cancer cells for cancer therapy. One potential drug target is isocitrate dehydrogenase 1 (IDH1), which is mutated in multiple human cancers. Here, we examine the role of mutant IDH1 in fully transformed cells with endogenous *IDH1* mutations. A selective R132H-IDH1 inhibitor (AGI-5198) identified through a high-throughput screen blocked, in a dose-dependent manner, the ability of the mutant enzyme (mIDH1) to produce *R*-2-hydroxyglutarate (*R*-2HG). Under conditions of near-complete *R*-2HG inhibition, the mIDH1 inhibitor induced demethylation of histone H3K9me3 and expression of genes associated with gliogenic differentiation. Blockade of mIDH1 impaired the growth of *IDH1*-mutant—but not *IDH1*–wild-type—glioma cells without appreciable changes in genome-wide DNA methylation. These data suggest that mIDH1 may promote glioma growth through mechanisms beyond its well-characterized epigenetic effects.

Somatic mutations in the metabolic enzyme isocitrate dehydrogenase (IDH) have recently been identified in multiple human cancers, including glioma (*1*, *2*), sarcoma (*3*, *4*), acute myeloid leukemia (*5*, *6*), and others. All mutations map to arginine residues in the catalytic pockets of IDH1 (R132) or IDH2 (R140 and R172) and confer on the enzymes a new activity: catalysis of alpha-ketoglutarate (2-OG) to the (*R*)-enantiomer of 2-hydroxyglutarate (*R*-2HG) (*7*, *8*). *R*-2HG is structurally similar to 2-OG and, due to its accumulation to millimolar concentrations in *IDH1*-mutant tumors, competitively inhibits 2-OG–dependent dioxygenases (*9*).

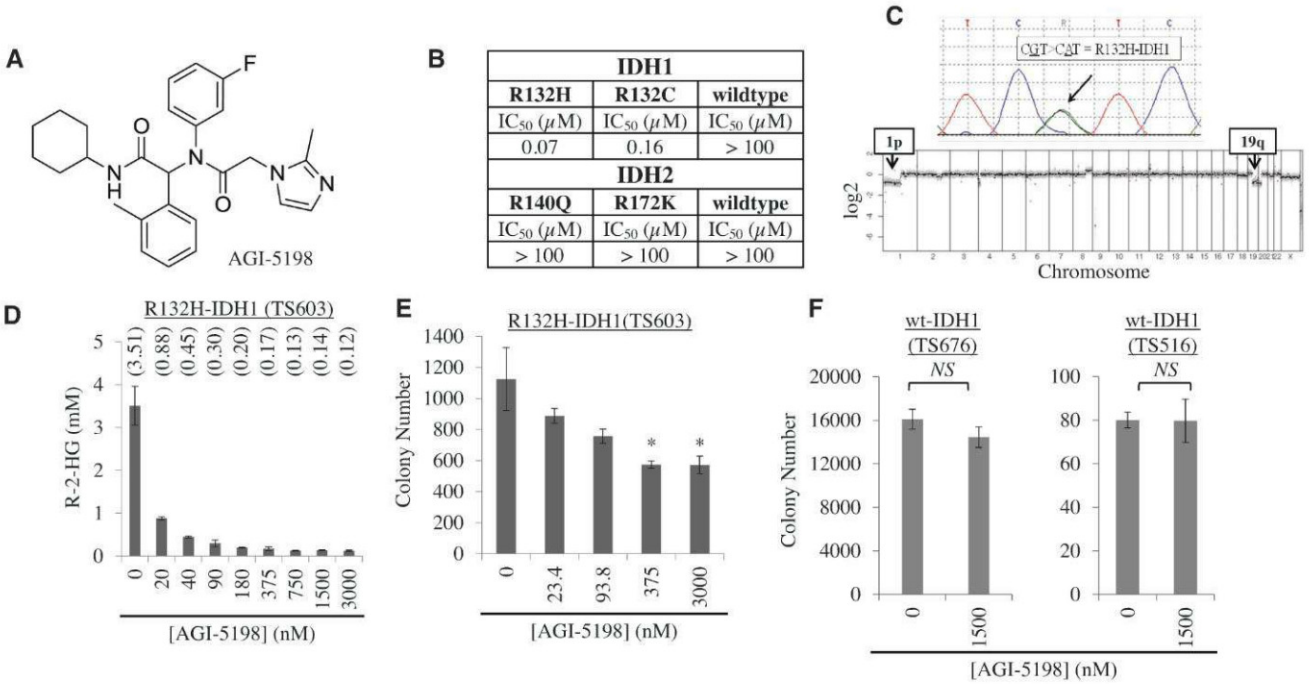
The mechanism by which mutant IDH1 contributes to the pathogenesis of human glioma remains incompletely understood. Mutations in *IDH1* are found in 50 to 80% of human low-grade (WHO grade II) glioma, a disease that progresses to fatal WHO grade III (anaplastic glioma) and WHO grade IV (glioblastoma) tumors over the course of 3 to 15 years. *IDH1* mutations appear to precede the occurrence of other mutations (*10*) and are associated with a distinctive gene-expression profile ("proneural" signature), DNA hypermethylation [CpG island methylator phenotype (CIMP)], and certain clinicopathological features (*11*–*13*). When ectopically expressed in immortalized human astrocytes, R132H-IDH1 promotes the growth of these cells in soft agar (*14*) and induces epigenetic alterations found in *IDH1*-mutant human gliomas (*15*, *16*). However, no tumor formation was observed when R132H-IDH1 was expressed from the endogenous *IDH1* locus in several cell types of the murine central nervous system (*17*).

To explore the role of mutant IDH1 in tumor maintenance, we used a compound that was identified in a high-throughput screen for compounds that inhibit the IDH1-R132H mutant homodimer (fig. S1 and supplementary materials) (*18*).

[1]Human Oncology and Pathogenesis Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA. [2]Department of Pharmacology, Weill-Cornell Graduate School of Biomedical Sciences, New York, NY 10021, USA. [3]Agios Pharmaceuticals, Cambridge, MA 02139, USA. [4]Department of Neurology, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA. [5]Department of Radiation Oncology, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA. [6]Department of Pathology, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA. [7]Department of Neurosurgery, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA. [8]Department of Molecular and Medical Pharmacology, Crump Institute for Molecular Imaging, University of California, Los Angeles, CA 90095, USA.

*These authors contributed equally to this work.
†This work is based on equal contributions from the laboratories of T.G.G., K.E.Y., and I.K.M.
‡Corresponding author. E-mail: katharine.yen@agios.com (K.E.Y.); mellingi@mskcc.org (I.K.M.)

This compound, subsequently referred to as AGI-5198 (Fig. 1A), potently inhibited mutant IDH1 [R132H-IDH1; half-maximal inhibitory concentration (IC50), 0.07 μM) but not wild-type IDH1 (IC50 > 100 μM) or any of the examined IDH2 isoforms (IC50 > 100 μM) (Fig. 1B). We observed no induction of nonspecific cell death at the highest examined concentration of AGI-5198 (20 μM).



IDH1		
R132H	R132C	wildtype
IC_{50} (μM)	IC_{50} (μM)	IC_{50} (μM)
0.07	0.16	> 100
IDH2		
R140Q	R172K	wildtype
IC_{50} (μM)	IC_{50} (μM)	IC_{50} (μM)
> 100	> 100	> 100

Fig. 1. An R132H-IDH1 inhibitor blocks *R*-2HG production and soft-agar growth of *IDH1*-mutant glioma cells. (**A**) Chemical structure of AGI-5198. (**B**) IC_{50} of AGI-5198 against different isoforms of IDH1 and IDH2, measured in vitro. (**C**) Sanger sequencing chromatogram (top) and comparative genomic hybridization profile array (bottom) of TS603 glioma cells. (**D**) AGI-5198 inhibits *R*-2HG production in *R132H-IDH1* mutant TS603 glioma cells. Cells were treated for 2 days with AGI-5198, and *R*-2HG was measured in cell pellets. *R*-2HG concentrations are indicated above each bar (in mM). Error bars, mean ± SEM of triplicates. (**E** and **F**) AGI-5198 impairs soft-agar colony formation of (E) *IDH1*-mutant TS603 glioma cells [*$P < 0.05$, one-way analysis of variance (ANOVA)] but not (F) *IDH1*–wild-type glioma cell lines (TS676 and TS516). Error bars, mean ± SEM of triplicates.

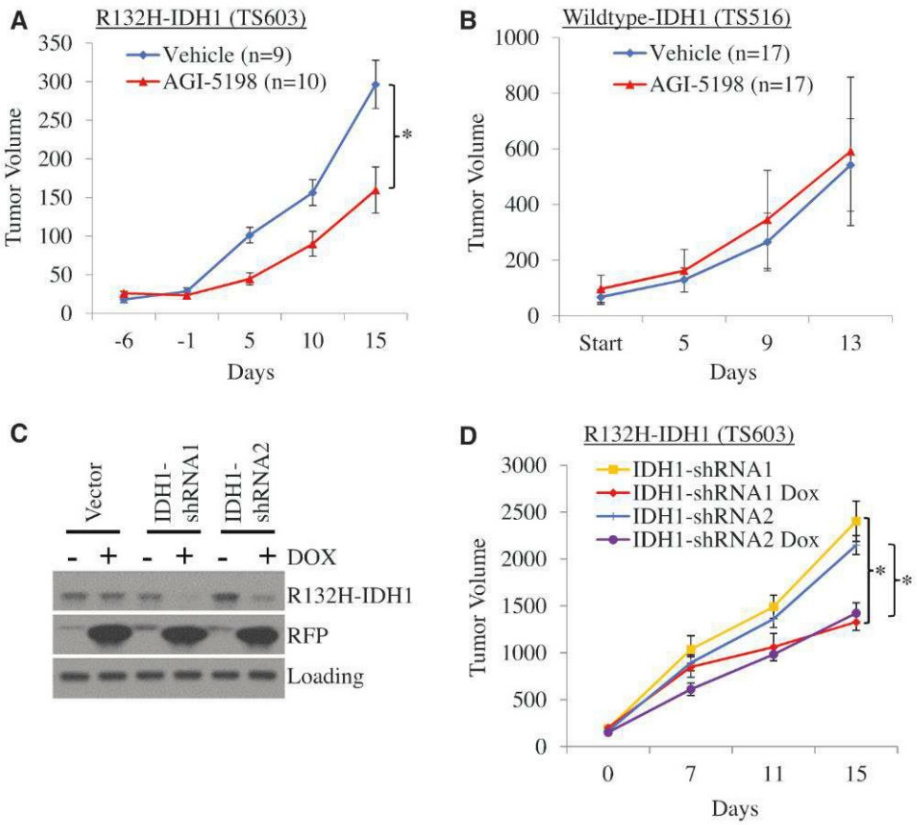


Fig. 2. AGI-5198 impairs growth of *IDH1*-mutant glioma xenografts in mice. (**A**) AGI-5198 impairs growth of *IDH1*±mutant TS603 glioma xenografts (*$P = 0.015$, two-tailed *t* test). Error bars, mean ± SEM. (**B**) AGI-5198 does not impair the growth of *IDH1*–wild-type TS516 glioma xenografts. Treatment with AGI-5198 (450 mg/kg per day via gavage) was started after subcutaneous xenografts were established. Error bars, mean ± SEM. (**C**) Inducible knockdown of IDH1 in *IDH1*-mutant glioma cells. Shown are Western blots from TS603 cells engineered to express either empty vector or IDH1-shRNAs from a doxycycline-inducible cassette and treated with doxycycline (2.5 μg/mL) for 6 days before lysis. (**D**) IDH1 knockdown impairs growth of *IDH1*-mutant glioma xenografts. Mice were randomized to vehicle versus doxycycline after subcutaneous tumors were established (*$P < 0.05$, two-tailed *t* test; $n = 15$ mice per cohort). Error bars, mean ± SEM.

We next explored the activity of AGI-5198 in TS603 glioma cells with an endogenous heterozygous *R132H-IDH1* mutation, the most common *IDH* mutation in glioma (*2*). TS603 cells were derived from a patient with anaplastic oligodendroglioma (WHO grade III) and harbor another pathognomomic lesion for this glioma subtype, namely co-deletion of the short arm of chromosome 1 (1p) and the long arm of chromosome 19 (19q) (*19*) (Fig. 1C). Measurements of *R*-2HG concentrations in pellets of TS603 glioma cells demonstrated dose-dependent inhibition of the mutant IDH1 enzyme by AGI-5198 (Fig. 1D). When added to TS603 glioma cells growing in soft agar, AGI-5198 inhibited colony formation by 40 to 60% (Fig. 1E). AGI-5198 did not impair colony formation of two patient-derived glioma lines that express only the wild-type *IDH1* allele (TS676 and TS516) (Fig. 1F), further supporting the selectivity of AGI-5198.

After exploratory pharmacokinetic studies in mice (fig. S2), we examined the effects of orally administered AGI-5198 on the growth of human glioma xenografts. When given daily to mice with established R132H-IDH1 glioma xenografts, AGI-5198 [450 mg per kg of weight (mg/kg) per os] caused 50 to 60% growth inhibition (Fig. 2A). Treatment was tolerated well with no signs of toxicity during 3 weeks of daily treatment (fig. S3). Tumors from AGI-5198–treated mice showed reduced staining with an antibody against the Ki-67 protein, a marker used for quantification of tumor cell proliferation in human brain tumors. In contrast, staining with an antibody against cleaved caspase-3 showed no differences between tumors from vehicle and AGI-5198–treated mice (fig. S4), suggesting that the growth-inhibitory effects of AGI-5198 were primarily due to impaired tumor cell proliferation rather than induction of apoptotic cell death. AGI-5198 did not affect the growth of *IDH1* wild-type glioma xenografts (Fig. 2B).

Given the likely prominent role of *R*-2HG in the pathogenesis of IDH-mutant human cancers, we investigated whether intratumoral depletion of this metabolite would have similar growth inhibitory effects on *R132H-IDH1*-mutant glioma cells as AGI-5198. We engineered TS603 sublines in which IDH1–short hairpin RNA (shRNA) targeting sequences were expressed from a doxycycline-inducible cassette. Doxycycline had no effect on IDH1 protein levels in cells expressing the vector control but depleted IDH1 protein levels by 60 to 80% in cells infected with *IDH1*-shRNA

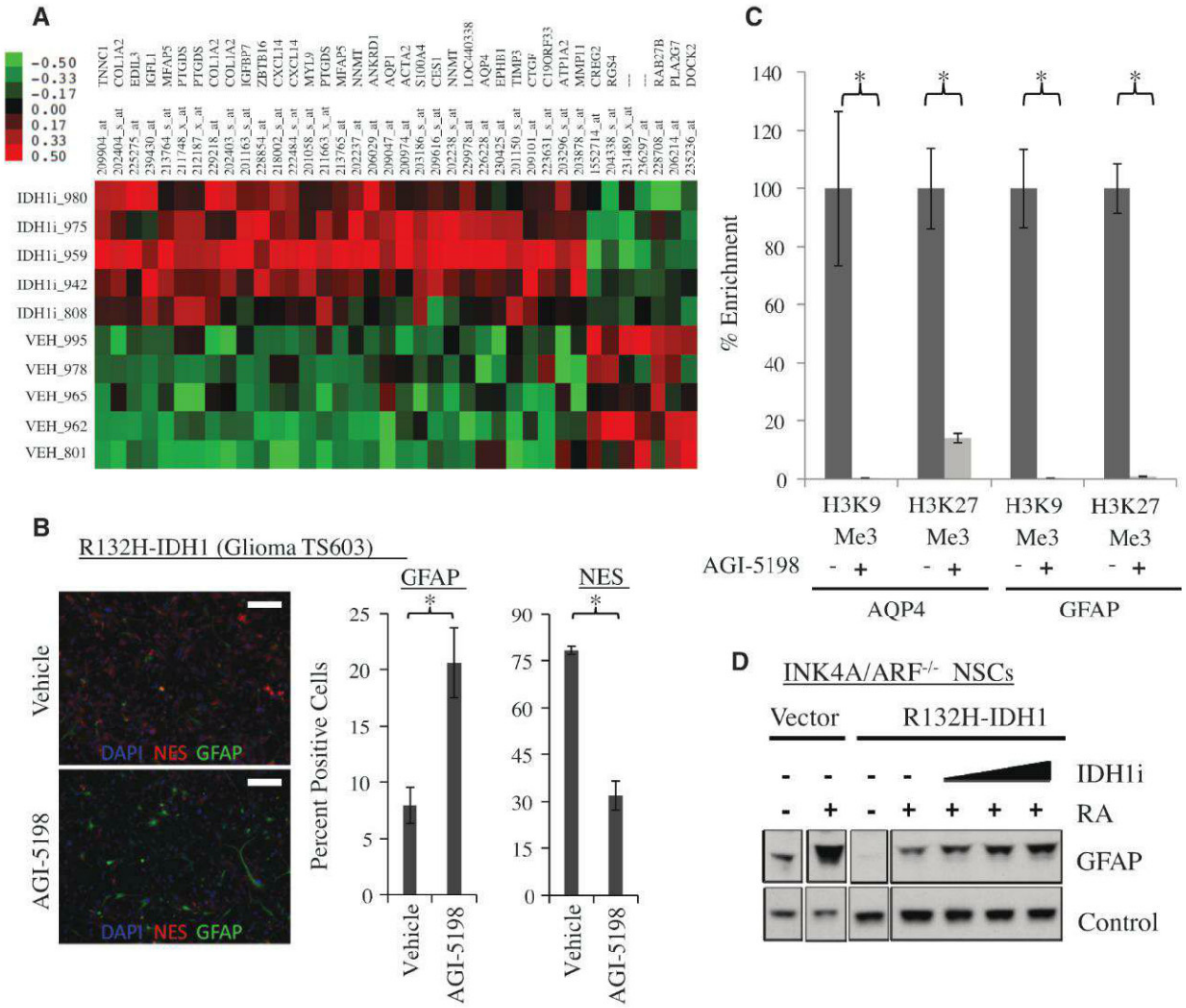


Fig. 3. AGI-5198 promotes astroglial differentiation in *R132H-IDH1* mutant cells. (**A**) Heat map of genes that are up- or down-regulated in TS603 glioma xenografts treated with AGI-5198 (more than twofold up or down). (**B**) Increased expression of GFAP (green) and decreased expression of NES (red) in TS603 cells treated in vitro with AGI-5198. Shown are immunofluorescence images of cells incubated in 1% fetal bovine serum (FBS) and 1 μM retinoic acid for 7 days in the presence of vehicle (top) or 1.5 μM AGI-5198 (bottom). Scale bar, 200 μM. The bar graph on the right represents a quantification of GFAP and NES staining (*$P < 0.05$, two-tailed *t* test). DAPI (4′,6-diamidino-2-phenylindole) staining in blue. Error bars, mean ± SEM of triplicates. (**C**) AGI-5198 promotes removal of repressive H3K9me3 and H3K27me3 marks at the GFAP and AQP4 promoters. Shown is ChIP (percent enrichment normalized to vehicle) of TS603 cells grown for 7 days in FBS (1%) and retinoic acid (1 μM) in the presence of vehicle or 1.5 μM AGI-5198 (*$P < 0.05$, two-tailed *t* test). Error bars, mean ± SEM for four repeats. (**D**) Blockade of mIDH1 restores the ability of *R132H-IDH1* mutant Ink4a/Arf$^{-/-}$ murine neuroprogenitor cells (NPCs) to express GFAP in response to retinoic acid. Shown is a Western blot of parental (vector) and R132H-IDH1 expressing Ink4a/Arf$^{-/-}$ NPCs treated with 1 μM retinoic acid (RA) in the presence of vehicle or mIDH1 inhibitor.

targeting sequences (Fig. 2C). We next injected these cells into the flanks of mice with severe combined immunodeficiency and, after establishment of subcutaneous tumors, randomized the mice to receive either regular chow or doxycycline-containing chow. As predicted from our experiments with AGI-5198, doxycycline impaired the growth of TS603 glioma cells expressing inducible IDH1-shRNAs in soft agar (fig. S5) and in vivo (Fig. 2D) but had no effect on the growth of tumors expressing the vector control (fig. S6). Immunohistochemistry (IHC) with a mutant-specific R132H-IDH1 antibody confirmed depletion of the mutant IDH1 protein in IDH1-shRNA tumors treated with doxycycline. This was associated with an 80 to 90% reduction in intratumoral *R*-2HG levels, similar to the levels observed in TS603 glioma xenografts treated with AGI-5198 (fig. S7). Knockdown of the IDH1 protein in *R132C-IDH1*-mutant HT1080 sarcoma cells similarly impaired the growth of these cells in vitro and in vivo (fig. S8).

To explore candidate molecular mechanisms of tumor growth inhibition by AGI-5198, we isolated RNA from the glioma xenografts in vehicle- and AGI-5198–treated mice, hybridized it to Affymetrix U133 plus 2.0 gene-expression arrays, and queried the RNA expression data for genes with a ≥2.0-fold increase or decrease in RNA expression in response to AGI-5198 (Fig. 3A) (table S1). The list of genes induced by AGI-5198 included several genes that are associated with astrocytic (aquaporin-4-AQP4; adenosine triphosphatase, Na^+/K^+ transporting, alpha 2 polypeptide-ATP1A2) (*20*, *21*) and oligodendrocytic differentiation [prostaglandin D2 synthase 21 kD (brain)-PTGDS] in the central nervous system (*21*). AGI-5198 also induced the expression of zinc finger and BTB domain–containing protein 16 (ZBTB16) [also known as promyelocytic leukemia zinc finger (PLZF)], a transcriptional repressor protein that is located on chromosome 11q23 and has been shown to promote glial differentiation in the central nervous system (*22*).

The gene-expression data suggested that treatment of *IDH1*-mutant glioma xenografts with AGI-5198 promotes a gene-expression program akin to gliogenic (i.e., astrocytic and oligodendrocytic) differentiation. To examine this question further, we treated TS603 glioma cells ex vivo with AGI-5198 and performed immunofluorescence for glial fibrillary acidic protein (GFAP) and nestin (NES) as markers for astrocytes and undifferentiated neuroprogenitor cells, respectively. Compared with vehicle, AGI-5198 treatment markedly increased the fraction of GFAP-positive cells and reduced the fraction of NES-positive cells (Fig. 3B). Chromatin-immunoprecipitation (ChIP) experiments showed that induction of the astrocytic markers GFAP and AQP4 in response to AGI-5198 was associated with marked reduction in repressive H3K9me3 and H3K27me3 marks on the promoters of these genes (Fig. 3C). Previous work showed that R132H-IDH1 expression in INK4A/ARF$^{-/-}$ murine neuroprogenitor cells (NPCs) restricts the ability of these cells to express GFAP in response to the differentiation cue retinoic acid (*23*). We investigated whether blockade of mutant IDH1 could restore this ability, and this was indeed the case (Fig. 3D). These results indicate that mIDH1 plays an active role in restricting cellular differentiation potential, and this defect is acutely reversible by blockade of the mutant enzyme.

In the developing central nervous system, gliogenic differentiation is regulated through changes in DNA and histone methylation (*24*). Mutant IDH1 can affect both epigenetic processes through *R*-2HG mediated suppression of TET (ten-eleven translocation) methyl cytosine hydroxylases and Jumonji-C domain histone demethylases (JHDMs). We therefore sought to define the epigenetic changes that were associated with the acute growth-inhibitory effects of AGI-5198 in vivo. We included a lower dose of AGI-5198 (150 mg/kg) in these experiments because of the differential sensitivity of 2-OG–dependent dioxygenases to competitive inhibition by *R*-2HG (*25*). We first examined intratumoral *R*-2HG concentrations after 2 weeks of treatment. Tumors from vehicle-treated mice showed *R*-2HG concentrations in the 4- to 6-mM range, similar to the concentrations of this metabolite in *IDH1*-mutant gliomas (*7*). Treatment of mice with AGI-5198 resulted in dose-dependent reduction of intratumoral *R*-2HG with partial *R*-2HG reduction at the 150 mg/kg dose (0.85 ± 0.22 mM) and near-complete reduction at the 450 mg/kg dose (0.13 ± 0.03 mM) (Fig. 4A).

We next examined whether acute pharmacological blockade of the mutant IDH1 enzyme reversed the CIMP, which is strongly associated with *IDH1*-mutant human gliomas (*12*). This CIMP phenotype is readily appreciated in untreated *R132H-IDH1* TS603 glioma xenografts using genome-wide DNA methylation arrays (Illumina Infinium Human Methylation 450 Arrays) by their increased number of highly methylated probes (β value > 0.7) (fig. S9) and by their unsupervised clustering with IDH-mutant human glioblastomas (fig. S10A) and intermediate-grade gliomas (fig. S10B). On a genome-wide scale,

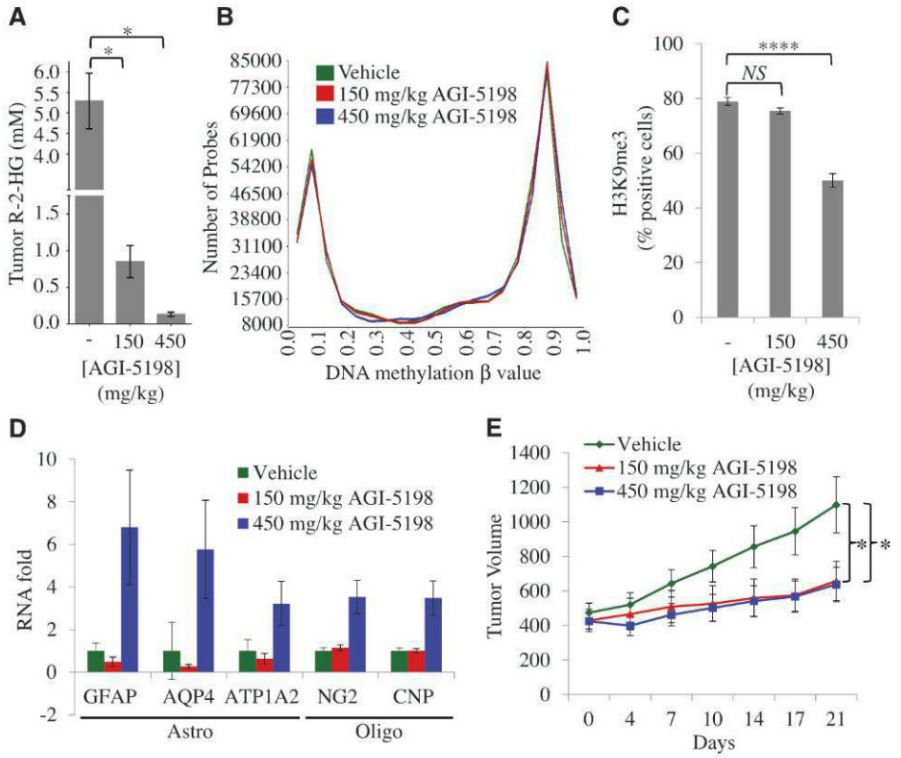


Fig. 4. Dose-dependent inhibition of histone methylation in *IDH1*-mutant gliomas after short-term treatment with AGI-5198. (**A**) Intratumoral concentrations of *R*-2HG in TS603 xenografts treated for 2 weeks with vehicle (n = 10 mice per cohort), 150 mg/kg AGI-5198 (n = 10 mice per cohort), or 450 mg/kg AGI-5198 (n = 8 mice per cohort) (*P < 0.05, two-tailed t test). Error bars, mean ± SEM. (**B**) Genome-wide distribution of DNA methylation in TS603 glioma xenografts treated for 2 weeks with vehicle, 150 mg/kg AGI-5198, or 450 mg/kg AGI-5198. (**C**) Effect of AGI-5198 on H3K9 trimethylation in TS603 glioma xenografts. Shown is the quantification of IHC results (****P < 0.00001) (see also fig. S11). Error bars, mean ± SEM of triplicates. (**D**) RNA levels of astrocytic (GFAP, AQP4, and ATP1A2) and oligodendrocytic (NG2 and CNP) differentiation markers in vehicle- versus AGI-5198–treated TS603 xenografts ($P = 4 \times 10^{-8}$, ANOVA). RNA expression was measured by RT-PCR. Error bars, mean ± SEM for 7 or 8 repeats (7 for vehicle, 8 for 150 mg/kg, and 7 for 450 mg/kg). (**E**) Tumor volumes of TS603 xenografts in mice treated with vehicle, 150 mg/kg AGI-5198, or 450 mg/kg AGI-5198 (*P < 0.05, two-tailed t test); n = 10 mice per cohort. Error bars, mean ± SEM.

we observed no statistically significant change in the distribution of β values between AGI-5198– and vehicle-treated tumors (Fig. 4B) (supplementary materials).

We next examined the kinetics of histone demethylation after inhibition of the mutant IDH1 enzyme. The histone demethylases JMJD2A and JMJD2C, which remove bi- and trimethyl marks from H3K9, are significantly more sensitive to inhibition by the *R*-2HG oncometabolite than other 2-OG–dependent oxygenases (*8*, *9*, *14*, *25*). Restoring their enzymatic activity in *IDH1*-mutant cancer cells would thus be expected to require near-complete inhibition of *R*-2HG production. Consistent with this prediction, tumors from the 450 mg/kg AGI-5198 cohort showed a marked decrease in H3K9me3 staining, but there was no decrease in H3K9me3 staining in tumors from the 150 mg/kg AGI-5198 cohort (Fig. 4C) (fig. S11). Of note, AGI-5198 did not decrease H3K9 trimethylation in *IDH1*–wild-type glioma xenografts (fig. S12A) or in normal astrocytes (fig. S12B), demonstrating that the effect of AGI-5198 on histone methylation was not only dose-dependent but also IDH1-mutant selective.

Because the inability to erase repressive H3K9 methylation can be sufficient to impair cellular differentiation of nontransformed cells (*16*), we examined the TS603 xenograft tumors for changes in the RNA expression of astrocytic (GFAP, AQP4, and ATP1A2) and oligodendrocytic (CNP and NG2) differentiation markers by real-time polymerase chain reaction (RT-PCR). Compared with vehicle-treated tumors, we observed an increase in the expression of astroglial differentiation genes only in tumors treated with 450 mg/kg AGI-5198 (Fig. 4D).

Despite its inability to reverse H3K9 trimethylation and induce gliogenic differentiation markers, the lower dose of AGI-5198 (150 mg/kg) resulted in a similar tumor growth inhibition as the higher dose of AGI-5198 (Fig. 4E) (fig. S13). Induction of the differentiation gene-expression program was thus associated with H3K9me3 demethylation but not required for tumor growth inhibition by AGI-5198, suggesting that mutant IDH regulated proliferation and differentiation in glioma through distinct effector programs with differential sensitivity, kinetic response, or reversibility to the *R*-2HG oncometabolite.

To identify pathways that are associated with the growth-inhibitory effects of AGI-5198, we ran Affymetrix RNA expression arrays. Many of the genes that showed significant changes in expression in both AGI-5198–treated cohorts relate to cardiovascular system development and tissue morphology (table S2) (supplementary materials). Interestingly, vascular abnormalities and disturbed collagen maturation were recently reported as the most prominent phenotype in mice with brain-specific expression of R132H-IDH1 (*17*).

When viewed on a genome-wide scale, the 150 mg/kg dose of AGI-5198, which was sufficient for maximal growth inhibition, showed only small effects on RNA expression patterns because 150 mg/kg-treated xenografts clustered with the untreated tumors (fig. S14). Furthermore, an integrated analysis of the DNA methylation and RNA expression data showed no correlation between changes in RNA expression and changes in DNA methylation (fig. S15). Together, these data suggest that mutant IDH1 may promote glioma growth through transcriptional and nontranscriptional mechanisms that are independent of its epigenetic effects.

In summary, we describe a tool compound (AGI-5198) that impairs the growth of *R132H-IDH1*-mutant, but not *IDH1* wild-type, glioma cells. This data demonstrates an important role of mutant IDH1 in tumor maintenance, in addition to its ability to promote transformation in certain cellular contexts (*14*, *26*). Effector pathways of mutant IDH remain incompletely understood and may differ between tumor types, reflecting clinical differences between these disorders. Although much attention has been directed toward TET-family methyl cytosine hydroxylases and Jumonji-C domain histone demethylases, the family of 2-OG–dependent dioxygenases includes more than 50 members with diverse functions in collagen maturation, hypoxic sensing, lipid biosynthesis/metabolism, and regulation of gene expression (*27*). Our study suggests that a broader investigation of the role of these enzymes in *IDH1*-mutant glioma may be warranted.

References and Notes

1. D. W. Parsons *et al.*, *Science* **321**, 1807 (2008).
2. H. Yan *et al.*, *N. Engl. J. Med.* **360**, 765 (2009).
3. M. F. Amary *et al.*, *J. Pathol.* **224**, 334 (2011).
4. D. R. Borger *et al.*, *Oncologist* **17**, 72 (2012).
5. P. Paschka *et al.*, *J. Clin. Oncol.* **28**, 3636 (2010).
6. E. R. Mardis *et al.*, *N. Engl. J. Med.* **361**, 1058 (2009).
7. L. Dang *et al.*, *Nature* **465**, 966 (2010).
8. M. E. Figueroa *et al.*, *Cancer Cell* **18**, 553 (2010).
9. W. Xu *et al.*, *Cancer Cell* **19**, 17 (2011).
10. T. Watanabe, S. Nobusawa, P. Kleihues, H. Ohgaki, *Am. J. Pathol.* **174**, 1149 (2009).
11. R. G. Verhaak *et al.*; Cancer Genome Atlas Research Network, *Cancer Cell* **17**, 98 (2010).
12. H. Noushmehr *et al.*; Cancer Genome Atlas Research Network, *Cancer Cell* **17**, 510 (2010).
13. A. Lai *et al.*, *J. Clin. Oncol.* **29**, 4482 (2011).
14. P. Koivunen *et al.*, *Nature* **483**, 484 (2012).
15. S. Turcan *et al.*, *Nature* **483**, 479 (2012).
16. C. Lu *et al.*, *Nature* **483**, 474 (2012).
17. M. Sasaki *et al.*, *Genes Dev.* **26**, 2038 (2012).
18. J. Popovici-Muller *et al.*, *ACS Medicinal Chem. Lett.* **3**, 850 (2012).
19. D. N. Louis *et al.*, *Acta Neuropathol.* **114**, 97 (2007).
20. J. D. Cahoy *et al.*, *J. Neurosci.* **28**, 264 (2008).
21. E. S. Lein *et al.*, *Nature* **445**, 168 (2007).
22. D. F. Sobieszczuk, A. Poliakov, Q. Xu, D. G. Wilkinson, *Genes Dev.* **24**, 206 (2010).
23. R. M. Bachoo *et al.*, *Cancer Cell* **1**, 269 (2002).
24. D. K. Ma *et al.*, *Nat. Neurosci.* **13**, 1338 (2010).
25. R. Chowdhury *et al.*, *EMBO Rep.* **12**, 463 (2011).
26. M. Sasaki *et al.*, *Nature* **488**, 656 (2012).
27. M. A. McDonough, C. Loenarz, R. Chowdhury, I. J. Clifton, C. J. Schofield, *Curr. Opin. Struct. Biol.* **20**, 659 (2010).

Acknowledgments: We thank W. G. Kaelin Jr., for sharing unpublished results. This work was supported by the National Brain Tumor Society (T.G.G., K.E.Y., and I.K.M.), the National Institutes of Health (1R01NS080944-01 and U54CA143798), the Leon Levy Foundation (I.K.M), and the James S. McDonnell Foundation (I.K.M.). C.G. was supported through an American Brain Tumor Association Basic Research Fellowship Award and the Lymphoma Research Foundation. T.G.G. is the recipient of a Research Scholar Award from the American Cancer Society. I.K.M. is the recipient of an Advanced Clinical Research Award in Glioma from the American Society of Clinical Oncology. Agios Pharmaceuticals has filed a patent application on AGI-5198. Microarray data are deposited in the Gene Expression Omnibus (GEO accession no. GSE45200).

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1236062/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S16
Tables S1 and S2
References

4 February 2013; accepted 18 March 2013
Published online 4 April 2013;
10.1126/science.1236062

Disparate Individual Fates Compose Robust CD8$^+$ T Cell Immunity

Veit R. Buchholz,[1]* Michael Flossdorf,[2,3]* Inge Hensel,[1] Lorenz Kretschmer,[1] Bianca Weissbrich,[1] Patricia Gräf,[1] Admar Verschoor,[1] Matthias Schiemann,[1,4] Thomas Höfer,[2,3]† Dirk H. Busch[1,4,5,6]†

A core feature of protective T cell responses to infection is the robust expansion and diversification of naïve antigen-specific T cell populations into short-lived effector and long-lived memory subsets. By means of in vivo fate mapping, we found a striking variability of immune responses derived from individual CD8$^+$ T cells and show that robust acute and recall immunity requires the initial recruitment of multiple precursors. Unbiased mathematical modeling identifies the random integration of multiple differentiation and division events as the driving force behind this variability. Within this probabilistic framework, cell fate is specified along a linear developmental path that progresses from slowly proliferating long-lived to rapidly expanding short-lived subsets. These data provide insights into how complex biological systems implement stochastic processes to guarantee robust outcomes.

Following Burnet's clonal selection theory (*1*), a single naïve lymphocyte should be the smallest unit from which an adaptive immune response can originate, providing both protection to acute infection and lasting memory to reinfection. Indeed, upon infection an individ-

ual CD8$^+$ T cell can generate a phenotypically and functionally diverse progeny that encompasses both short- and long-lived subsets (*2*, *3*). Thus, individual T cells possess a stunning capacity for diversification (*4*). However, how this capacity is channeled to generate reproducible immune responses tailored to a specific invading pathogen is unknown. Two mechanisms to achieve this goal can be envisioned: (i) a single–T cell–derived response follows a defined expansion and diversification pattern, thus guaranteeing a robust relationship of input and response on the single-cell level; or (ii) this relationship only becomes robust on the level of multiple T cells responding to antigen, as a population integral of variable individual fates. Furthermore, it remains unresolved in which developmental order long- and short-lived subsets arise; different models propose either subset as a predecessor of the other (*5*, *6*) or suggest asymmetric cell division as key to their simultaneous generation (*7*).

To provide a comprehensive understanding of these issues for CD8$^+$ T cells, we have combined an adoptive transfer approach that allows mapping the fate of multiple single-cell–derived progenies in vivo with mathematical modeling. We crossed OT-I T cell receptor (TCR)–transgenic mice, whose CD8$^+$ T cells are specific for the SIINFEKL peptide from chicken ovalbumin (OVA) presented on major histocompatibility complex I, to C57BL/6 mice expressing the congenic markers CD45.1 and/or CD90.1 in a homo- or heterozygous fashion. Thereby, we created a matrix of eight TCR-identical OT-I donor lines, distinguishable from one another and from C57BL/6 recipients by their congenic phenotypes (Fig. 1A). Antibody labeling allowed us to discriminate matrix donors from one another and from the CD45.2/CD45.2–CD90.2/CD90.2 recipient C57BL/6 mice (fig. S1). Successive high-purity single-cell sorting of CD8$^+$CD44^{low} cells (fig. S2) displaying homogenous TCR expression levels and naïve phenotype (fig. S3) enabled the assembly of eight individual OT-I matrix cells for subsequent adoptive transfer (fig. S4). To assess the expansion and diversification of single CD8$^+$ T cell–derived progenies into effector and memory subsets, we infected recipient mice with recombinant *Listeria monocytogenes*–expressing OVA (*L.m.*-OVA). In our hands and as previously shown (*8*), the differentiation and expansion patterns of low numbers of transferred OT-I T cells closely mirrored those of endogenous SIINFEKL epitope-specific T cells (fig. S5). Simultaneous adoptive transfer

[1]Institute for Medical Microbiology, Immunology and Hygiene, Technische Universität München (TUM), Munich 81675, Germany. [2]Division of Theoretical Systems Biology, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany. [3]BioQuant Center, University of Heidelberg, Heidelberg 69120, Germany. [4]Clinical Cooperation Groups "Antigen-specific Immunotherapy" and "Immune-Monitoring," Helmholtz Center Munich (Neuherberg), TUM, Munich 81675, Germany. [5]Focus Group "Clinical Cell Processing and Purification," Institute for Advanced Study, TUM, Munich 81675, Germany. [6]National Centre for Infection Research (DZIF), Munich 81675, Germany.

*These authors contributed equally to this work.
†Corresponding authors. E-mail: dirk.busch@mikrobio.med.tum.de (D.H.B.); t.hoefer@dkfz-heidelberg.de (T.H.)

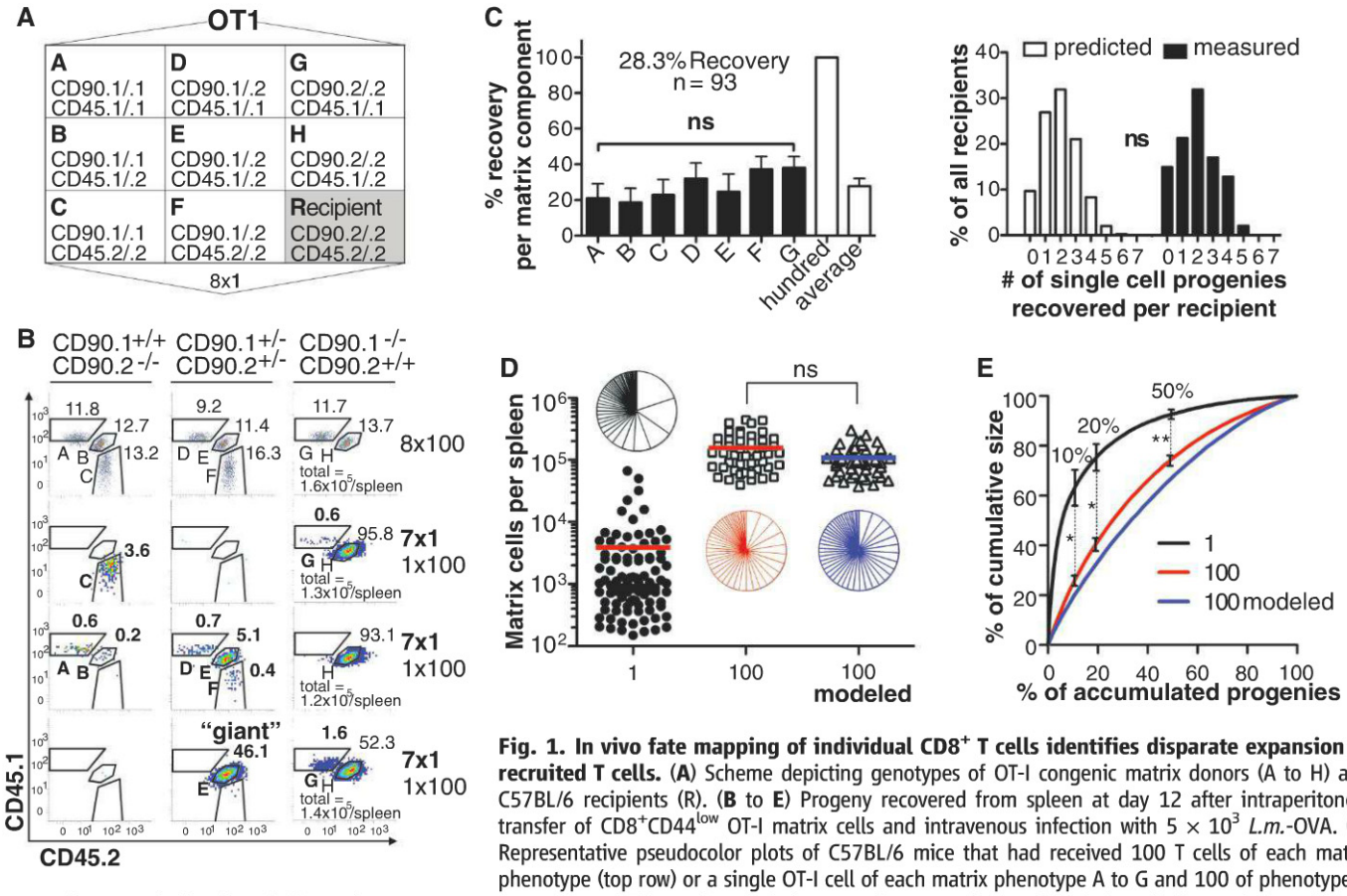


Fig. 1. In vivo fate mapping of individual CD8$^+$ T cells identifies disparate expansion of recruited T cells. (**A**) Scheme depicting genotypes of OT-I congenic matrix donors (A to H) and C57BL/6 recipients (R). (**B** to **E**) Progeny recovered from spleen at day 12 after intraperitoneal transfer of CD8$^+$CD44^{low} OT-I matrix cells and intravenous infection with 5×10^3 *L.m.*-OVA. (B) Representative pseudocolor plots of C57BL/6 mice that had received 100 T cells of each matrix phenotype (top row) or a single OT-I cell of each matrix phenotype A to G and 100 of phenotype H (bottom three rows). Numbers indicate the percentage that a given OT-I matrix population contributes to the combined size of all populations in one recipient. (C) Percent in which progeny could be recovered from single or 100 transferred OT-I cells of matrix phenotype A to G or H, respectively (left). Predicted versus measured distribution of number of single-cell–derived progenies recovered per recipient (right); *n* is the number of single-cell progenies recovered from a total of 47 C57BL/6 recipients; Bars depict mean; error bars indicate SEM. (D) Absolute number of descendants in spleen recovered per transferred single cell (full circles) ($n = 93$), 100 cells (white squares) ($n = 47$), or generated by mathematical modeling of 100-cell transfers from single-cell data (white triangles). Each triangle consists of 28 progenies randomly drawn from the single-cell–derived data set. Bars depict mean. Wheel diagrams depict the relative contribution of each progeny to the cumulative size of all. (E) Cumulative size of single-cell–derived (black line), 100-cell–derived (red line), and mathematically modeled progenies (blue line) plotted against the percentage of accumulated progenies. Significant differences of cumulative size are indicated for accumulating the 10, 20, and 50% largest progenies; error bars indicate SEM.

of 100 OT-I T cells of each matrix phenotype ("8×100" transfer) led to comparable expansion and subset diversification of all eight matrix populations (Fig. 1B, top row, and fig. S6).

To compare the progenies of individual cells and cell populations, every C57BL/6 recipient received seven single-matrix OT-I cells (matrix components A to G) and one congenic population consisting of 100 matrix OT-I cells (matrix component H) in a combined "7×1 plus 1×100" transfer (Fig. 1B). Independent of matrix phenotype, descendants were recovered from ~28% of transferred single cells (Fig. 1C). This indicates that a transferred 100-cell matrix population contributes on average 28 detectable single-cell–derived progenies to the overall response. Indeed, the means of single- and 100-cell–derived progeny sizes scaled according to this prediction (Fig. 1D). However, size distribution of single-cell–derived progenies was remarkably broad, ranging over three orders of magnitude (Fig. 1D). By day 12 after infection, most single cells had generated numbers of descendants well below the mean of ~4000 cells per spleen ("dwarfs"), whereas very few expanded massively, generating up to 70,000 descendants ("giants") (Fig. 1, B, bottom row, and D). The number of descendants recovered from the spleen tightly correlated to a progeny's overall expansion in spleen, lung, and lymph nodes (fig. S7). Highly disparate single-cell–derived expansion occurred within individual recipients (Fig. 1B, bottom) and could neither be accounted for by inter-individual infection variability nor variable expression of additional TCRs on transferred OT-I T cells (figs. S8 and S9). We could recreate in silico the much narrower distribution of 100-cell–derived populations through repetitively drawing 28 random samples from our single-cell–derived data set (Fig. 1, D and E). The skewed size distribution of single-cell progenies leads to 5% of naïve precursors generating over 50% of all descendants (Fig. 1E). Similarly skewed expansion patterns were observed before the contraction phase, at day 8 after infection (fig. S10), making it unlikely that cell death substantially shapes disparate single-cell–derived progeny sizes. Single-cell–derived responses remained equally skewed when individual precursors were transferred without a 100-cell control population (fig. S11), excluding any unphysiological competitive pressure as the cause of skewness. A different immunization using systemic infection with Vaccinia virus–expressing OVA (Vaccinia-OVA) yielded congruent results (fig. S12). Thus, within a monoclonal population of T cells the bulk of infection-driven expansion is based on the proliferative activity of a small fraction of recruited precursors.

To probe the physiological consequences of heterogeneous single-cell–derived expansion, we analyzed surface markers associated with memory development. Strikingly, most single-cell–derived progenies expressed significantly higher levels of memory markers CD62L (*9*) and CD27 (*10*) than expected from the expression patterns of 100-cell–derived populations (Fig. 2, A to C, and figs. S13 and S14). However, the rare single-cell progenies that had expanded to giants were strongly biased for $CD27^{-}CD62L^{-}$ effector phenotype (Fig. 2D, top left). This inverse relation between proliferation and memory phenotype (Fig. 2, D to F, and figs. S13 and S14) accounts for the different distribution of memory markers in single-cell progenies and 100-cell–derived populations, as shown by composing 100-cell transfer data from single-cell progenies in silico (Fig. 2, B and C).

Next, we studied the production of cytokines. Autochthonous interleukin-2 (IL-2) production during primary infection signifies memory or memory precursor cells (*9*, *11*, *12*). Coproduction of interferon-γ (IFN-γ) is characteristic of multifunctional cells, the presence of which is thought to predict vaccination success (*13*). Similar to the expression pattern of surface memory markers, most single-cell–derived progenies showed a higher percentage of IL-2 and IFN-γ producers than expected from population analysis (Fig. 3, A to C). Whereas expression of IL-2 was essentially absent and IFN-γ production low in giant progenies with a dominant effector phenotype, the opposite was true for the abundant dwarfs (Fig. 3, D to F). These also expressed less T-box transcription factor expressed in T cells (T-bet) and showed a trend toward higher eomesodermin (Eomes) expression than that of their giant counterparts (Fig. 3G). The association of Eomes with memory and T-bet with terminal effector fate (*14*–*16*) was further confirmed by their opposing correlation patterns in relation to expression of CD62L, CD27, and short-lived effector cell marker Killer-cell-lectin-like-receptor-G1 (KLRG1) (Fig. 3H) (*17*). Thus, heterogeneous single–T cell–derived expansion was reproducibly accompanied by a relative reduction of memory precursor characteristics—on the level of phenotype, function, and transcription factor expression.

To alter stimulus strength, we delayed transfer of single T cells to day 3 after infection. As expected, this reduced mean response size (*18*). However, size variability between single-cell progenies remained virtually unaffected (fig. S15).

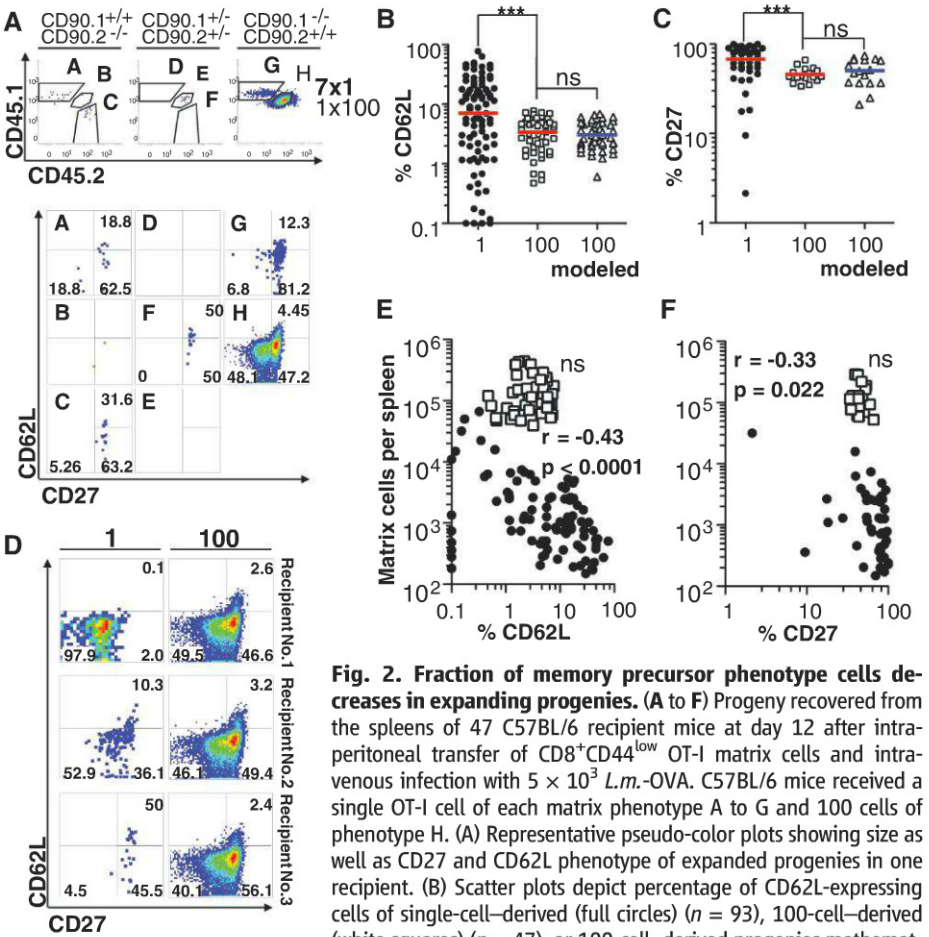


Fig. 2. Fraction of memory precursor phenotype cells decreases in expanding progenies. (**A** to **F**) Progeny recovered from the spleens of 47 C57BL/6 recipient mice at day 12 after intraperitoneal transfer of $CD8^{+}CD44^{low}$ OT-I matrix cells and intravenous infection with 5×10^{3} *L.m.*-OVA. C57BL/6 mice received a single OT-I cell of each matrix phenotype A to G and 100 cells of phenotype H. (A) Representative pseudo-color plots showing size as well as CD27 and CD62L phenotype of expanded progenies in one recipient. (B) Scatter plots depict percentage of CD62L-expressing cells of single-cell–derived (full circles) (n = 93), 100-cell–derived (white squares) (n = 47), or 100-cell–derived progenies mathematically modeled from the single-cell–derived data set (white triangles). Bars indicate median. (C) As in (B), but for percentage of CD27-expressing cells. (D) Representative pseudo-color plots showing CD27 and CD62L expression of large to small single-cell–derived progenies together with phenotype of 100-cell–derived progenies in the same recipients. (E) Scatter plots depict correlation of size and percentage of CD62L-expressing cells of single-cell–derived (full circles) and 100-cell–derived progenies (white squares). (F) As in (E), but for percentage of CD27-expressing cells.

Thus, global alteration of immune response parameters proved unfit to reach mechanistic insight into how variability develops on the single-cell level. To provide more insight, we asked whether single-cell progenies at the peak of the immune response contain a "footprint" of their individual differentiation and proliferation histories. For unbiased computational analysis of this question, we constructed all possible diversification pathways of a single naïve cell into three derived subsets: central memory precursor (TCMp; $CD27^+CD62L^+$), effector memory precursor (TEMp; $CD27^+CD62L^-$), and effector cells (TEF; $CD27^-CD62L^-$). Allowing for both unidirectional and reversible differentiation steps and subset-specific proliferation rates, we obtained 304 distinct pathways of T cell diversification from the general scheme (Fig. 4A) and translated these into stochastic dynamic models (fig. S16) (*19–21*). Only two models provided a good fit to the experimental data, with all parameter values being identifiable within narrow bounds (Fig. 4B and fig. S17). All remaining 302 models fit the data poorly (Fig. 4C and fig. S18). The two fitting models predict a linear framework of differentiation, following the pathway naïve→TCMp→TEMp→TEF [in Fig. 4B, model 1, ~10% of cells develop directly from naïve to TEMp, which could be due to a fraction of naïve cells dividing asymmetrically (fig. S19); model 2 fits the data nearly as well without this transition]. The proliferation rates increase along this pathway, with TCMp cells proliferating substantially slower than TEMp and TEF cells (figs. S17 and S19). The linear diversification model correctly reproduces the inverse correlation between memory marker (CD62L and CD27) abundance and size of single-cell progenies at day 8 after infection (Fig. 4D and fig. S20). To further test the model, we simulated and subsequently measured these sizes and correlations at day 6 after infection (Fig. 4D and fig. S21) as well as the proliferation rates at days 3 and 4 after infection (fig. S22), finding agreement between model and experimental data. Relying only on single-cell progeny data from day 8 after infection, the model correctly predicted the phenotypic composition of an expanding T cell population over days 1 to 8 after infection (Fig. 4E and fig. S22). Thus, our data imply a linear framework of $CD8^+$ T cell differentiation, with slowly proliferating memory precursors giving rise to rapidly expanding effector cells.

Our model identifies the initial events in each single-cell progeny's developmental history as the drivers of variability. As long as the number of daughter cells is very low (up to ~10 descendants), the stochastic timing of proliferation and differentiation events (mediated through intrinsic and/or extrinsic cues) will cause individual progenies to differ strongly from one another. These differences are stably propagated when larger cell numbers are reached (Fig. 4F). In particular, the early transition of single-cell–derived descendants into higher proliferative states can give rise to giants dominated by rapidly expanding TEMp and TEF cells (fig. S23). Overall, our data imply that robustness of $CD8^+$ T cell immune responses requires participation of ~50 naïve progenitors (fig. S24).

Single-cell progeny size correlated strongly with the size of TEMp and TEF compartments and only weakly with the absolute number of TCMp cells (Fig. 4G), despite the existence of a strong inverse correlation between the relative fraction of TCMp cells and progeny size (Fig. 4D, left). This implies that a progeny's primary size is a poor predictor of its expansion potential upon secondary challenge. To examine this

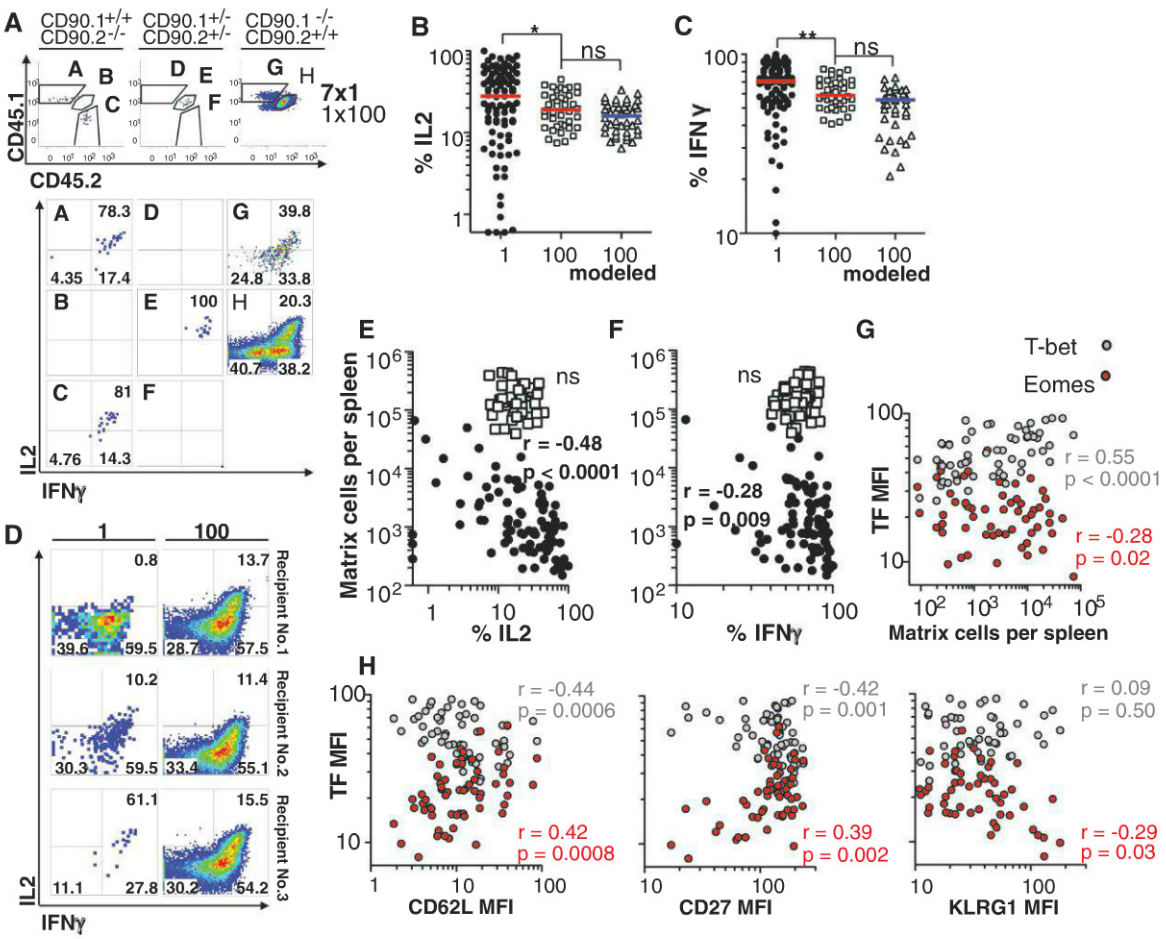


Fig. 3. Fraction of cells with memory precursor function and respective transcription factor profile decreases in expanding progenies. (**A** to **F**) As in Fig. 2, A to F, but for percentage of IL-2– and IFNγ-expressing cells. (**G** and **H**) Scatter plots depict mean fluorescence intensity (MFI) of Eomes (red dots) and T-bet (gray dots) correlated to (G) size or (H) MFI of CD62L, CD27, and KLRG1 of single-cell–derived progenies.

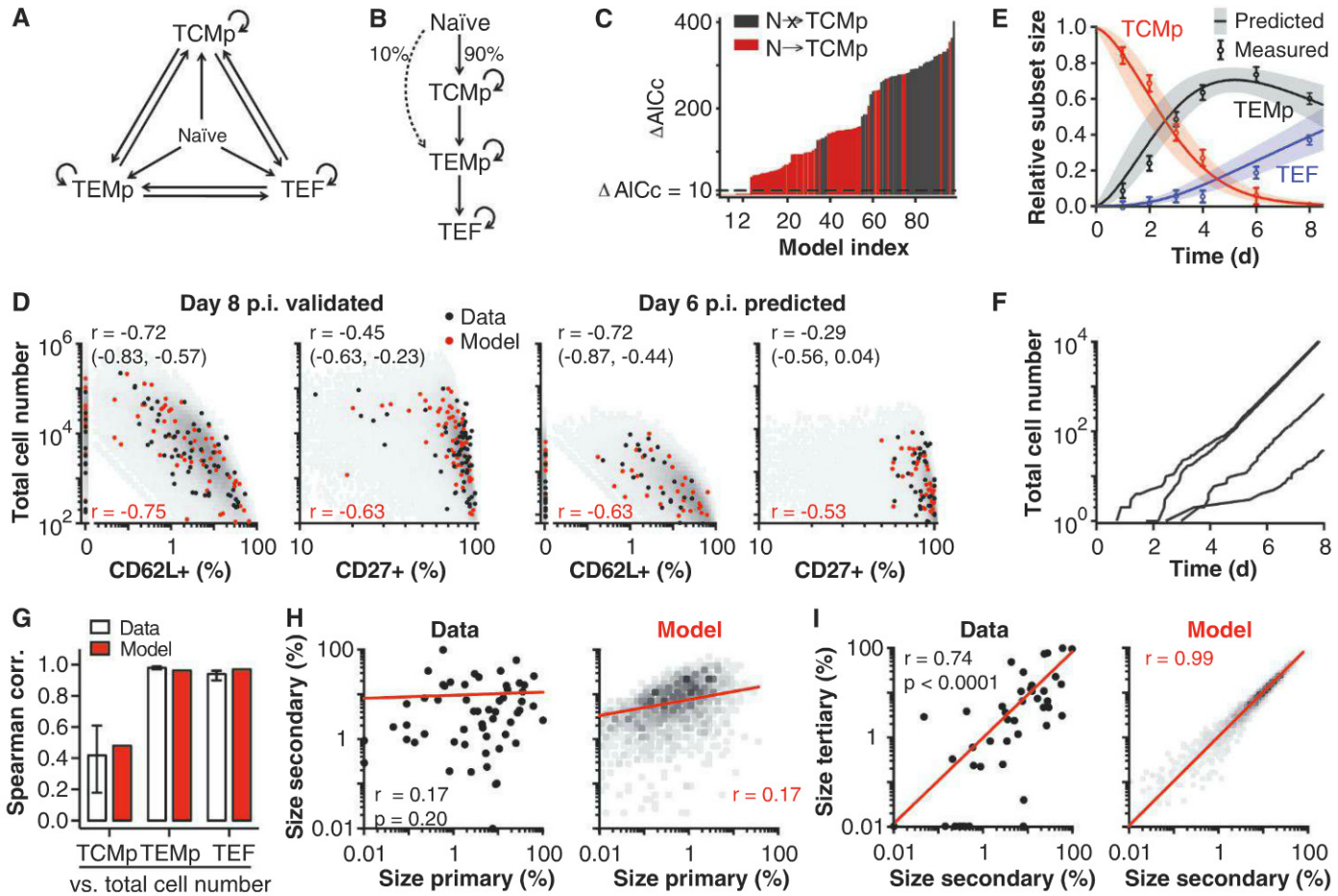


Fig. 4. Stochastic progression along a linear differentiation pathway underlies CD8⁺ T cell diversification. (**A**) Model scheme allowing for all possible diversification pathways from naïve to TCMp, TEMp, and TEF subsets, with subset-specific differentiation (straight arrows) and proliferation (curved arrows) rates. (**B**) Unique model structure that accounts for the experimental data. (**C**) Model ranking with corrected Akaike information criterion; models having ΔAICc > 10 are essentially unsupported by the data. Red bars indicate models including a naïve→TCMp differentiation step; gray bars indicate other models. (**D**) As in Fig. 2, E and F, but for days 8 and 6 after infection; measured single-cell–derived progenies (black dots), equal number of stochastic simulations (red dots), and probability density (gray scale) of naïve→TCMp→TEMp→TEF model. p.i., post-infection. (**E**) Relative subset sizes at days 1 to 8 after infection predicted with the model exclusively from day 8 single-cell progeny data (filled regions indicate 95% prediction bands) and measured (circles) after transfer of 10,000 (days 1 to 4) or 100 OTI cells (days 6 and 8). Error bars indicate SEM and are based on pooled variance for days 1 to 6. (**F**) Four simulations of single-cell progeny expansion. (**G**) Correlation of absolute TCMp, TEMp, and TEF number versus total cell number; error bars, 95% confidence interval. (**H** and **I**) Data (left) and model prediction (right) for relation of (H) primary versus secondary and (I) secondary versus tertiary response size (largest progeny = 100%, others scaled accordingly); log-log regression (red line); experiment performed as in Fig. 2, but progenies recovered from peripheral blood at day 8 after primary (day 8 after infection), day 6 after secondary (day 42 after infection), and day 6 after tertiary infection (day 235 after infection) with 5×10^3 *L.m.*-OVA, 1×10^5 *L.m.*-OVA, and 1×10^8 MVA-OVA, respectively.

hypothesis, we monitored the peak size of single-cell–derived progenies during primary and secondary immune responses (fig. S25). Indeed, peak size during secondary challenge was nearly uncorrelated to primary progeny size, allowing some "dwarfs" to expand up to 200-fold in relation to their primary size (Fig. 4H). However, when animals were subjected to heterologous tertiary challenge with Modified Vaccinia Ankara–expressing OVA (MVA-OVA), we observed a strong correlation to secondary response size (Fig. 4I). Both data sets are in accord with the naïve→TCMp→TEMp→TEF model when recall expansion is assumed to originate mainly from TCMp populations (Fig. 4, H and I). Together with similar findings gathered by means of genetically labeling individual T cells (*22*), these observations argue that memory content is not predicted by a single-cell progeny's primary, effector-dominated expansion, yet there is a robust and predictable expansion pattern of recall responses originating from populations of multiple memory cells.

Collectively, our data suggest that the early developmental histories of single-cell–derived progenies shape their disparate fates, which compose robust immunity only together, as distinct variations on a shared developmental theme. Our findings are reminiscent of stochastic precursor progeny dynamics and proliferative hierarchies, recently described as essential for stem cell–based maintenance of tissue homeostasis (*23*, *24*). By shedding light on how the descendants of individual T cells compose complex immune responses, we provide a framework for the strategic design of future vaccination and adoptive immunotherapy approaches.

References and Notes

1. F. M. Burnet, *The Clonal Selection Theory of Acquired Immunity* (Cambridge Univ. Press, Cambridge, UK, 1959).
2. C. Stemberger *et al.*, *Immunity* **27**, 985 (2007).
3. C. Gerlach *et al.*, *J. Exp. Med.* **207**, 1235 (2010).
4. S. L. Reiner, F. Sallusto, A. Lanzavecchia, *Science* **317**, 622 (2007).
5. E. J. Wherry *et al.*, *Nat. Immunol.* **4**, 225 (2003).
6. A. V. Gett, F. Sallusto, A. Lanzavecchia, J. Geginat, *Nat. Immunol.* **4**, 355 (2003).
7. J. T. Chang *et al.*, *Science* **315**, 1687 (2007).
8. V. P. Badovinac, J. S. Haring, J. T. Harty, *Immunity* **26**, 827 (2007).
9. F. Sallusto, D. Lenig, R. Förster, M. Lipp, A. Lanzavecchia, *Nature* **401**, 708 (1999).
10. H. Hikono *et al.*, *J. Exp. Med.* **204**, 1625 (2007).

11. M. A. Williams, A. J. Tyznik, M. J. Bevan, *Nature* **441**, 890 (2006).
12. S. Feau, R. Arens, S. Togher, S. P. Schoenberger, *Nat. Immunol.* **12**, 908 (2011).
13. R. A. Seder, P. A. Darrah, M. Roederer, *Nat. Rev. Immunol.* **8**, 247 (2008).
14. A. Banerjee *et al.*, *J. Immunol.* **185**, 4988 (2010).
15. X. Zhou *et al.*, *Immunity* **33**, 229 (2010).
16. N. S. Joshi *et al.*, *Immunity* **27**, 281 (2007).
17. D. Voehringer *et al.*, *J. Immunol.* **167**, 4838 (2001).
18. W. N. D'Souza, S. M. Hedrick, *J. Immunol.* **177**, 777 (2006).
19. J. E. Till, E. A. McCulloch, L. Siminovitch, *Proc. Natl. Acad. Sci. U.S.A.* **51**, 29 (1964).
20. K. R. Duffy *et al.*, *Science* **335**, 338 (2012).
21. Materials and methods are available as supplementary materials on *Science* Online.
22. C. Gerlach *et al.*, *Science* **340**, 635 (2013).
23. B. D. Simons, H. Clevers, *Cell* **145**, 851 (2011).
24. G. Mascré *et al.*, *Nature* **489**, 257 (2012).

Acknowledgments: The authors thank L. Henkel for cell sorting; A. Muschaweckh and I. Drexler for providing Vaccinia-OVA and MVA-OVA; and K. Kerksiek, H. Wagner, C. Stemberger, F. Anderl, S. Dreher, and T. G. Grünewald for helpful discussion and/or critically reading the manuscript. The authors declare no conflict of interest. The data presented in this manuscript are tabulated in the main paper and in the supplementary materials. This work was supported by the SFB TR 36 (TP-B10/13, TP-Z1), the SFB 1054 (TP-B09), the SFB 914 (TP-B04), the Initiative and Networking Fund of the Helmholtz Association within the Helmholtz Alliance on Immunotherapy of Cancer, EU-FP7 SYBILLA (201106), BMBF ForSysPartner (0315267E), and by the National Science Foundation under grant NSF PHY11-25915.

Supplementary Materials
www.sciencemag.org/cgi/content/full/science.1235454/DC1
Materials and Methods
Figs. S1 to S27
References (*25–28*)

22 January 2013; accepted 7 March 2013
Published online 14 March 2013;
10.1126/science.1235454

Heterogeneous Differentiation Patterns of Individual CD8$^+$ T Cells

Carmen Gerlach,[1]*† Jan C. Rohr,[1,2]* Leïla Perié,[1,3] Nienke van Rooij,[1] Jeroen W. J. van Heijst,[1]‡ Arno Velds,[4] Jos Urbanus,[1] Shalin H. Naik,[1]§ Heinz Jacobs,[5] Joost B. Beltman,[1,3] Rob J. de Boer,[3] Ton N. M. Schumacher[1]‖

Upon infection, antigen-specific CD8$^+$ T lymphocyte responses display a highly reproducible pattern of expansion and contraction that is thought to reflect a uniform behavior of individual cells. We tracked the progeny of individual mouse CD8$^+$ T cells by in vivo lineage tracing and demonstrated that, even for T cells bearing identical T cell receptors, both clonal expansion and differentiation patterns are heterogeneous. As a consequence, individual naïve T lymphocytes contributed differentially to short- and long-term protection, as revealed by participation of their progeny during primary versus recall infections. The discordance in fate of individual naïve T cells argues against asymmetric division as a singular driver of CD8$^+$ T cell heterogeneity and demonstrates that reproducibility of CD8$^+$ T cell responses is achieved through population averaging.

The murine naïve T cell repertoire contains about 100 to 1000 CD8$^+$ T cells specific for a given antigen (*1*, *2*), which upon antigen-recognition proliferate to produce up to 10^7 progeny (*2–4*). After pathogen clearance, the numbers of antigen-specific T cells decline by 90 to 95%, leaving behind a population of memory T cells (*5*). This characteristic T cell response kinetic is accompanied by differentiation into functionally distinct subsets (*6*, *7*).

Although the patterns of T cell differentiation and response kinetics are highly reproducible for a given infection, it is unclear how this reproducibility is controlled. Single-cell tracing studies have unambiguously shown that individual naïve T cells are able to yield both effector and memory cells (*8*, *9*). However, the relative contribution of individual naïve T cells to the effector and memory pools has not been determined. Reproducibility of T cell responses could be controlled at the single-cell level, with each naïve T cell producing the same types and amount of progeny, and asymmetric division of T cells (*10*) would provide a mechanism to ensure such equal representation. Alternatively, individual T cell families (i.e., one naïve T cell and its progeny) may show distinct sizes or phenotypes, with reproducibility manifesting itself only at the population level. Both scenarios predict very different mechanisms behind protection to renewed infection. In the former, a defined fraction of each family provides long-term memory; in the latter, some T cell families could primarily convey short-term protection, whereas others could mostly yield memory cells.

We aimed to distinguish between these two potential mechanisms for T cell reactivity during primary and renewed infections by in vivo lineage tracing of individual naïve CD8$^+$ T cells. Combination of a previously established DNA barcode–based lineage tracing technology (*9*, *11–13*) with second-generation sequencing allowed quantification of individual T cell families with substantial accuracy [quantification and correction procedures are described in fig. S1 and (*14*)]. To track individual T cells in a system in which differences in T cell receptor (TCR) affinity do not influence cell fate, we generated naïve TCR-transgenic OT-I T cells [which recognize the SIINFEKL (E, Glu; F, Phe; I, Ile; K, Lys; L, Leu; N, Asn; and S, Ser) peptide presented on H-2K^b] that each carry a unique DNA barcode (*9*). Mice transferred with physiological numbers of these cells were then infected with *Listeria monocytogenes* expressing SIINFEKL (LM-OVA), and barcode sequencing was used to quantify the progeny of individual naïve T cells.

Analysis of the contribution of individual T cells at the peak of the immune response revealed the number of progeny produced by individual T cells to be highly variable (Fig. 1A, controls in fig. S2) and, on average, the "dominant" naïve T cell produced 400-fold more offspring than the median naïve T cell within the same animal (Fig. 1B). As a consequence, about 60% of the total OT-I T cell pool was derived from only 5% of naïve OT-I cells (Fig. 1C). Although ~50% of OT-I families consisted of fewer than 200 daughter cells, these minifamilies had formed as a consequence of TCR triggering rather than bystander or homeostatic proliferation (fig. S3A). The numerical dominance of a small number of T cell families was unrelated to rearrangement of the endogenous TCR loci, because LM-OVA–induced responses of OT-I *Rag2*$^{-/-}$ T cells were also biased toward few families (fig. S3, B to D). Furthermore, family-size disparity was independent of the number of responding OT-I families (range evaluated:17 to 874 families per animal, fig. S3E), arguing against competition between the transferred cells as a confounding factor.

To evaluate whether these results were influenced by the retroviral integration sites of barcodes, we developed a transgenic mouse strain (BCM) that allows DNA barcode generation in vivo at a defined genomic locus (fig. S4A). After LM-OVA infection, OT-I T cells harboring such endogenously generated barcodes also displayed a marked variability in clonal output (fig. S4, B and C). Thus, T cell fate is also disparate when cells are tagged at a specific genomic site.

Strong disparity in the number of progeny of different naïve T cells constitutes a robust feature of T cell responses, observed upon infection with different doses of LM-OVA (Fig. 1D), in high- and low-affinity TCR-antigen interactions (Fig. 1E) and upon pulmonary influenza infection

[1]Division of Immunology, Netherlands Cancer Institute, Amsterdam, Netherlands. [2]Centre of Chronic Immunodeficiency (CCI), University Medical Centre Freiburg and University of Freiburg, Freiburg, Germany. [3]Theoretical Biology and Bioinformatics, Utrecht University, Utrecht, Netherlands. [4]Central Genomics Facility, Netherlands Cancer Institute, Amsterdam, Netherlands. [5]Division of Biological Stress Responses, Netherlands Cancer Institute, Amsterdam, Netherlands.

*These authors contributed equally to this work.
†Present address: Department of Microbiology and Immunobiology, Harvard Medical School, Boston, MA 02115, USA.
‡Present address: Immunology Program, Sloan-Kettering Institute, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA.
§Present address: Division of Immunology, The Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville Victoria 3052, Australia.
‖Corresponding author. E-mail: t.schumacher@nki.nl

(fig. S5). Kinetic analyses revealed that clonal disparity was already apparent at day 5 and remained similarly pronounced during the following days (Fig. 2, A to D, and fig. S6). An appealing explanation would be that naïve T cells enter their first division at different time points postinfection and hence differ in the time window of clonal expansion. However, mathematical modeling based on carboxyfluorescein diacetate succinimidyl ester (CFSE) data suggests that the time of first division cannot be the main decisive factor (fig. S7). Furthermore, strong variation in family size was also observed when naïve T cells were transferred 48 hours after infection (fig. S8), indicating that individual "late-comer" T cells (*15*) also show disparate behavior.

Collectively, the above experiments reveal that the numerical output of individual T cells is highly variable. As a consequence, the magnitude of T cell responses only becomes reproducible when the output of some 50 T cell families is combined (Fig. 2E). This heterogeneity appeared consistent with two distinct scenarios for the formation of T cell memory. In a first model, some naïve T cells are imprinted to expand more strongly during the primary expansion, and the resulting families would remain similarly dominant during recall infection or become increasingly prominent should this imprinting still exist (fig. S9A). In an opposite model, family-size disparity reflects a division of labor, in which some naïve T cells mainly contribute offspring for primary responses whereas others primarily seed T cell memory (fig. S9B).

The second model predicts that different T cell families show substantial variation with respect to phenotype and, more directly, with respect to contribution to recall responses. To evaluate this, we first separated barcode-labeled OT-I T cells after LM-OVA infection on the basis of expression of three different markers of T differentiation state, KLRG-1, CD27, and CD62L. Analysis of the distribution of individual T cell families over these subsets revealed that individual T cell families showed pronounced variation in phenotype, as revealed by the fraction of cells expressing KLRG-1, CD27, or CD62L (Fig. 3A and fig. S10). Thus, not only the magnitude but also the phenotypic properties of antigen-specific T cell responses are controlled by averaging disparate behaviors of individual T cell families.

We next analyzed whether T cell family size was correlated to the differentiation state of its members. This revealed a strong inverse correlation between T cell family size and the percentage of CD62L-expressing cells (Fig. 3B), indicating that clonal expansion and loss of CD62L expression may share a common regulator, as previously proposed (*16*). In contrast, the fraction of KLRG-1–positive T cells showed a weaker (positive) correlation with family size. Furthermore, the fraction of CD27-positive cells and family size were only marginally correlated, both at days 7 and 11 (Fig. 3B), indicating differential regulation of family size, KLRG-1 and CD27 expression.

To directly evaluate whether the offspring of individual T cells contributes differently to early protection and long-term immunity, we quantified T cell families longitudinally within the same animal. First, comparison of clonal dominance patterns during primary response and resting memory phase showed that the extent of contraction (i.e., the fraction of cells that survived into the memory phase) was essentially independent of family size (fig. S11). Subsequently, we quantified family sizes during three consecutive infections (Fig. 4A). Unsupervised clustering showed that barcode distribution was much more alike between secondary and tertiary infections than

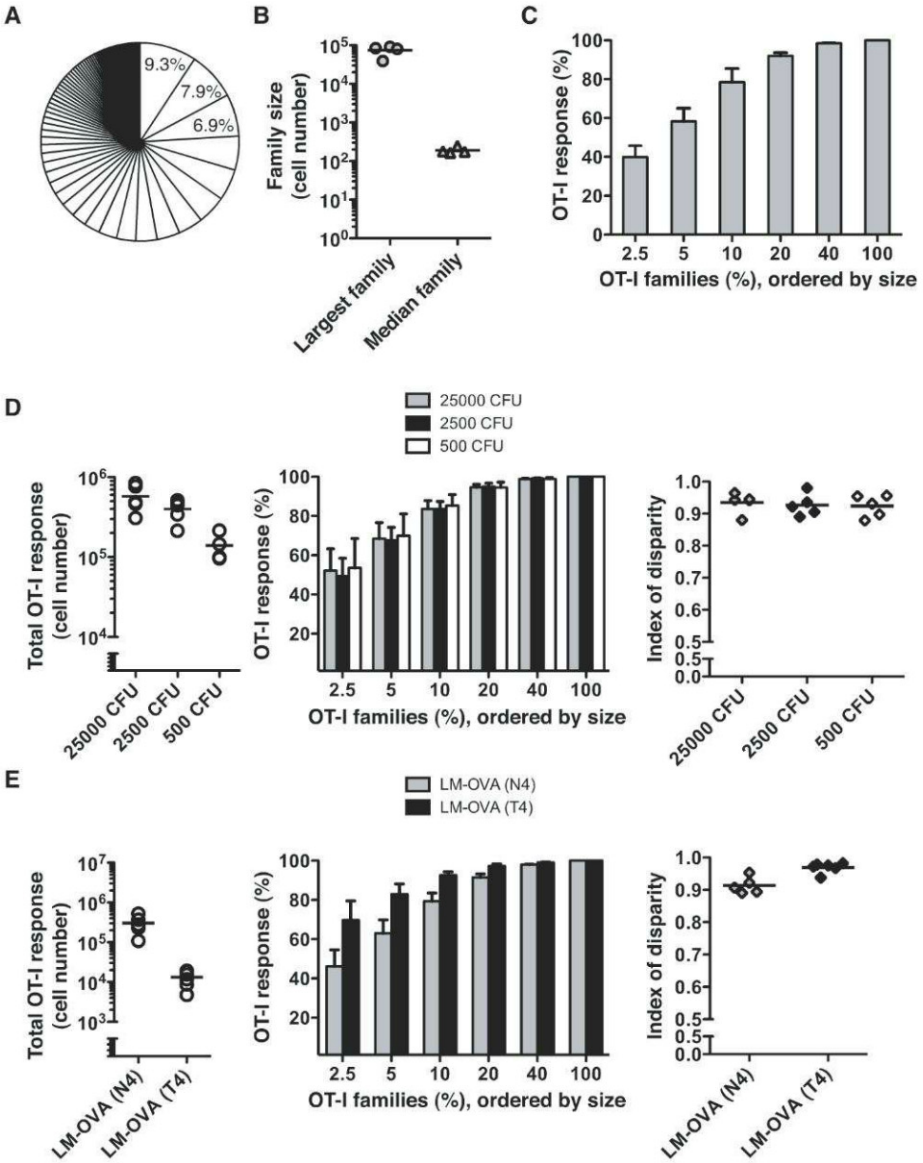


Fig. 1. Single-cell output disparity within CD8$^+$ T cell responses. Recipients of 800 naïve, barcode-labeled OT-I T cells were infected with LM-OVA. At day 7 postinfection, OT-I T cell family sizes in spleen and lymph nodes (LNs) were quantified. (**A**) Average size of the largest to smallest OT-I families, represented as fraction of the overall OT-I T cell response derived from, on average, 228 naïve T cells. (**B**) Number of OT-I cells constituting the largest and median OT-I families. Symbols indicate individual mice; bar, average. (**C**) Average (+SD) fraction of overall OT-I response formed by the indicated fractions of OT-I families, ordered by dominance. Data from one experiment (four mice), representative of 13. (**D** and **E**) Disparity between T cell families in recipients receiving (D) 1000 or (E) 640 barcode-labeled OT-I cells and infected with (D) indicated doses of LM-OVA and (E) LM-OVA (N4) or an LM-OVA variant expressing SIITFEKL (T4). (Left) OT-I T cell response size (number of GFP$^+$ cells). CFU, colony-forming units. (Middle) Fraction (mean + SD) of OT-I T cell response formed by the indicated fractions of OT-I families, ordered by dominance. (Right) Index of disparity for OT-I family sizes. Upper limit of the index of disparity is 1, occurring when one family dominates. The lower limit is 0, occurring when all families occupy an equal fraction. Data from one experiment per experimental setting (≥4 mice per group), representative of two each. (A to D) Lineage-tracing quality controls are shown in fig. S2.

between any of these and the primary T cell response (Fig. 4B). Comparison of the size of individual T cell families during primary and secondary infections revealed that many of the dominant T cell families during the primary immune response were present at substantially lower frequency during secondary infection and vice versa (Fig. 4C). This highly variable participation in secondary responses reflected an intrinsic difference between T cell families in their ability to reexpand, rather than stochastic variation, because the pattern of clonal dominance during the secondary response was very well preserved upon tertiary infection (Fig. 4, D and E). In line with this, when memory T cells from one mouse were transferred into two recipient mice, the patterns of clonal dominance were similar in these paired recipients upon subsequent infection and, as expected (Fig. 4C), distinct from the pattern observed during the primary T cell response within the matched donor mouse (Fig. 4, F to H, and fig. S12).

Our experiments demonstrate a division of labor between T cell families in immune protection to renewed infection (fig. S9B). Because T cell families show disparate behavior along multiple axes that are only partially correlated (e.g., clone size versus fraction CD27$^+$; Fig. 3), we consider it unlikely that variation in one single factor is responsible for this heterogeneity. Furthermore, our data imply that at least some T cell fates must be heritably imprinted before clonal expansion becomes substantial. Specifically, the observation that T cell families of similar size can consist of cells that are either largely devoid of CD27 or essentially all CD27-positive implies an early fixation of this fate.

Collectively, our data and the data from Buchholz *et al.* (*17*) demonstrate that the behavior of individual T cells varies markedly with respect to clonal expansion, differentiation, and recall capacity. This variability may be due to both intrinsic stochastic variation, as has been observed for B cells in vitro (*18*), and to variation in extrinsic signals, such as local differences in

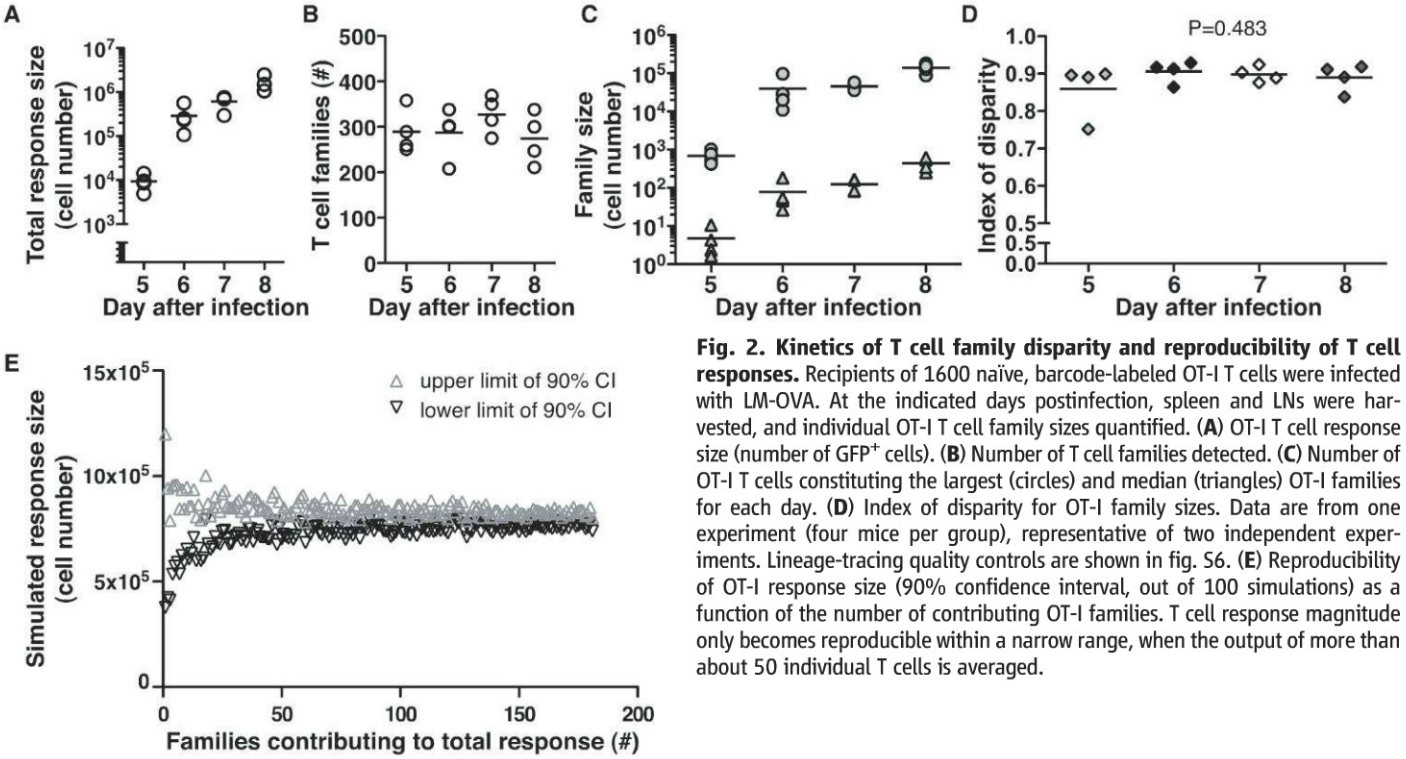


Fig. 2. Kinetics of T cell family disparity and reproducibility of T cell responses. Recipients of 1600 naïve, barcode-labeled OT-I T cells were infected with LM-OVA. At the indicated days postinfection, spleen and LNs were harvested, and individual OT-I T cell family sizes quantified. (**A**) OT-I T cell response size (number of GFP$^+$ cells). (**B**) Number of T cell families detected. (**C**) Number of OT-I T cells constituting the largest (circles) and median (triangles) OT-I families for each day. (**D**) Index of disparity for OT-I family sizes. Data are from one experiment (four mice per group), representative of two independent experiments. Lineage-tracing quality controls are shown in fig. S6. (**E**) Reproducibility of OT-I response size (90% confidence interval, out of 100 simulations) as a function of the number of contributing OT-I families. T cell response magnitude only becomes reproducible within a narrow range, when the output of more than about 50 individual T cells is averaged.

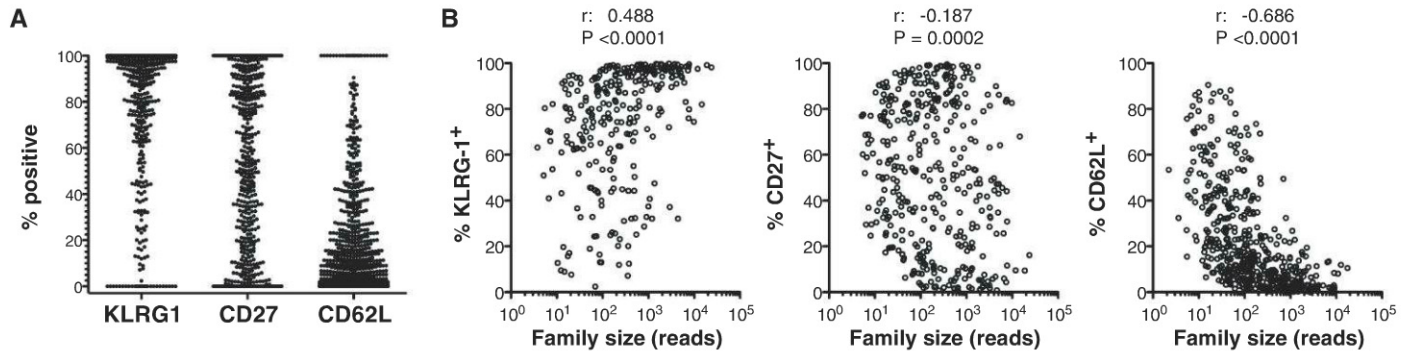


Fig. 3. Individual T cells yield phenotypically distinct progenies. Recipients of 500 to 700 naïve, barcode-labeled OT-I T cells were infected with LM-OVA. At day 11 postinfection, spleen and LNs were harvested. OT-I T cells were isolated on the basis of expression of KLRG1, CD27, or CD62L by cell sorting, and T cell family sizes in the indicated phenotypic subsets were quantified. Plots depict pooled data from four mice. (**A**) For each T cell family, the fraction of members expressing the indicated marker is depicted. (**B**) For each T cell family, the fraction of members expressing the indicated marker is depicted as a function of the relative size of this family. Lineage-tracing quality controls are provided in fig. S10. The Spearman correlation *r* was calculated for all families contributing to the marker-positive and -negative populations. Data are from one experiment (four mice) per marker, representative of two independent experiments per marker.

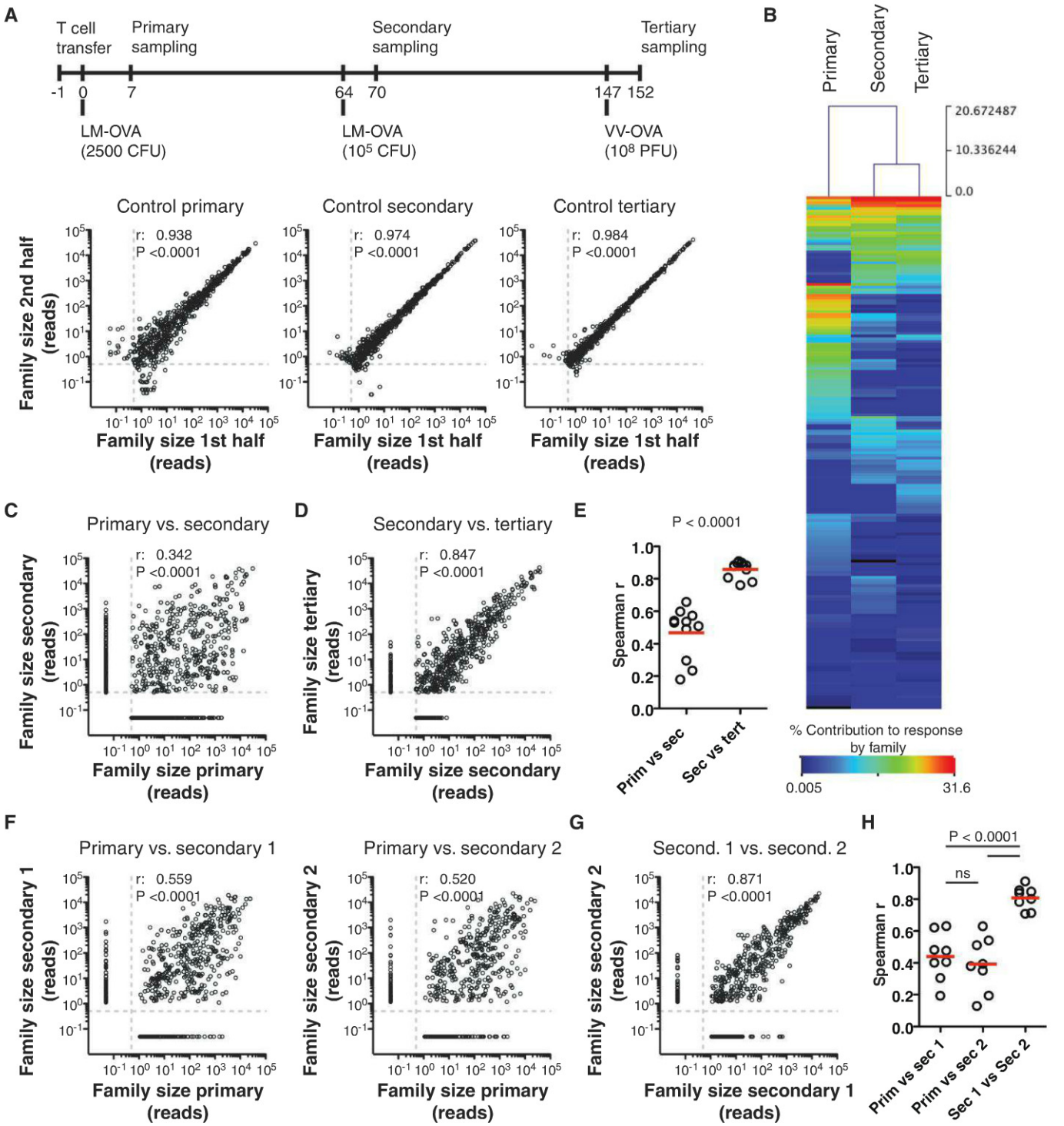


Fig. 4. Division of labor between antigen-specific T cell families. (**A**) Experimental setup and lineage-tracing controls for (B) to (E). VV-OVA abbreviates recombinant Vaccinia virus encoding SIINFEKL. (**B**) Euclidean distance heatmap representing contribution of individual naïve T cells (rows) to the primary-tertiary response (columns). Data from one representative mouse of two experiments with at least five mice per experiment. (**C**) Comparison of family size (barcode frequencies) between primary and secondary responses. (**D**) As (C), but for secondary and tertiary responses. Data for (C) to (D) pooled from five mice of one experiment, representative of two, with each mouse receiving 500 to 600 cells. (**E**) Spearman correlation *r* between family sizes in indicated samples calculated for individual mice. Pooled data from two experiments. Significance of differences analyzed by paired *t* test. (F to H) Recipients of barcode-labeled OT-I T cells were infected with LM-OVA. At day 59, $CD8^+$ T cells were transferred into secondary recipients, followed by LM-OVA infection (fig. S12B). OT-I family sizes were quantified at day 8 after primary and day 6 after secondary infection. (**F**) Comparison of family size between primary and secondary responses. (**G**) Comparison between matched secondary recipients. Data for (F) and (G) pooled from four mice of one experiment, representative of two, with each mouse receiving 500 to 600 cells. (**H**) Spearman correlation as in (E). Pooled data from two experiments. Depicted barcode frequencies are normalized to 10^5 per half sample. Read counts of 0 set to 0.05 for log-scale visualization. Gray lines, cut-off for barcode detection. Spearman correlations were calculated for all families contributing to both depicted responses. Lineage-tracing controls are in fig. S12.

antigen-presenting cells or cytokines. In either case, the current observations exclude models in which each naïve T cell yields progeny with the same distribution of cells of either short- or long-term potential. Thus, although our data do not evaluate the role of asymmetric division as a mechanism to generate daughter cells with different fates, they do show that asymmetric division by itself cannot explain the disparity between individual T cell families that is experimentally observed. Rather, a strong variation between families in the expansion of proximal and distal daughter cells needs to be invoked to arrive at the heterogeneity observed here. Lastly, the observation of strong heterogeneity at the single-cell level indicates that T cell responses are made up of "averages," a behavior reminiscent of recent analyses of stem-cell renewal (*19*). Thus, although the differentiation and expansion of the combined T cell population follow a uniform course, the fate of individual naïve T cells is unpredictable.

References and Notes

1. J. J. Obar, K. M. Khanna, L. Lefrançois, *Immunity* **28**, 859 (2008).
2. J. N. Blattman *et al.*, *J. Exp. Med.* **195**, 657 (2002).
3. K. Murali-Krishna *et al.*, *Immunity* **8**, 177 (1998).
4. E. A. Butz, M. J. Bevan, *Immunity* **8**, 167 (1998).
5. M. A. Williams, M. J. Bevan, *Annu. Rev. Immunol.* **25**, 171 (2007).
6. S. Sarkar *et al.*, *J. Exp. Med.* **205**, 625 (2008).
7. J. J. Obar, L. Lefrançois, *Int. Immunol.* **22**, 619 (2010).
8. C. Stemberger *et al.*, *Immunity* **27**, 985 (2007).
9. C. Gerlach *et al.*, *J. Exp. Med.* **207**, 1235 (2010).
10. J. T. Chang *et al.*, *Science* **315**, 1687 (2007).
11. J. W. J. van Heijst *et al.*, *Science* **325**, 1265 (2009).
12. T. N. M. Schumacher, C. Gerlach, J. W. J. van Heijst, *Nat. Rev. Immunol.* **10**, 621 (2010).
13. K. Schepers *et al.*, *J. Exp. Med.* **205**, 2309 (2008).
14. Materials and methods are available as supplementary materials on *Science* Online.
15. W. N. D'Souza, S. M. Hedrick, *J. Immunol.* **177**, 777 (2006).
16. T. E. Schlub *et al.*, *Eur. J. Immunol.* **39**, 67 (2009).
17. V. R. Buchholz *et al.*, *Science* **340**, 630 (2013).
18. K. R. Duffy *et al.*, *Science* **335**, 338 (2012).
19. B. D. Simons, H. Clevers, *Cell* **145**, 851 (2011).

Acknowledgments: The authors thank A. Pfauth and F. van Diepen for cell sorting; E. Borg for technical assistance; D. Zehn for provision of LM-OVA (T4); S. Harkema for helpful discussions; and G. Bendle, M. C. Wolkers, and E. A. Moseman for critically reading the manuscript. The data presented in this manuscript are tabulated in the main paper and in the supplementary materials. The authors declare no conflict of interest. This research was funded by European Research Council AdG Life-his-T to T.N.M.S., a Walter-Hitzig fellowship from the CCI (BMBF 01EO0803), a Deutsche Forschungsgemeinschaft (DFG RO 4120/1-1) research fellowship to J.C.R., a Marie Curie Intra European Fellowship to L.P., a Nederlandse Organisatie voor Wetenschappelijk Onderzoek Veni grant (916.86.080) to J.B.B., and a Leukemia and Lymphoma Society Career Development Fellowship to S.H.N.

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1235487/DC1
Materials and Methods
Supplementary Text
Fig. S1
Table S1
References (*20*–*23*)

22 January 2013; accepted 7 March 2013
Published online 14 March 2013;
10.1126/science.1235487

Neural Decoding of Visual Imagery During Sleep

T. Horikawa,[1,2] M. Tamaki,[1]* Y. Miyawaki,[3,1]† Y. Kamitani[1,2]‡

Visual imagery during sleep has long been a topic of persistent speculation, but its private nature has hampered objective analysis. Here we present a neural decoding approach in which machine-learning models predict the contents of visual imagery during the sleep-onset period, given measured brain activity, by discovering links between human functional magnetic resonance imaging patterns and verbal reports with the assistance of lexical and image databases. Decoding models trained on stimulus-induced brain activity in visual cortical areas showed accurate classification, detection, and identification of contents. Our findings demonstrate that specific visual experience during sleep is represented by brain activity patterns shared by stimulus perception, providing a means to uncover subjective contents of dreaming using objective neural measurement.

Dreaming is a subjective experience during sleep often accompanied by vivid visual contents. Previous research has attempted to link physiological states with dreaming (*1*–*3*), but none has demonstrated how specific visual contents are represented in brain activity. The advent of machine-learning–based analysis allows for the decoding of stimulus- and task-induced brain activity patterns to reveal visual contents (*4*–*9*). We extended this approach to the decoding of spontaneous brain activity during sleep (Fig. 1A). Although dreaming has often been associated with the rapid-eye movement (REM) sleep stage, recent studies have demonstrated that dreaming is dissociable from REM sleep and can be experienced during non-REM periods (*10*). We focused on visual imagery (hallucination) experienced during the sleep-onset (hypnagogic) period (sleep stage 1 or 2) (*11*, *12*) because it allowed us to collect many observations by repeating awakenings and recording participants' verbal reports of visual experience. Reports at awakenings in sleep-onset and REM periods share general features such as frequency, length, and contents, while differing in several aspects, including the affective component (*13*–*15*). We analyzed verbal reports using a lexical database to create systematic labels for visual contents. We hypothesized that contents of visual imagery during sleep are represented at least partly by visual cortical activity patterns shared by stimulus representation. Thus, we trained decoders on brain activity induced by natural images from Web image databases.

Three people participated in the functional magnetic resonance imaging (fMRI) sleep experiments (Fig. 1A), in which they were woken when an electroencephalogram signature was detected (*16*) (fig. S1), and they were asked to give a verbal report freely describing their visual experience before awakening [table S1; duration, 34 ± 19 s (mean ± SD)]. We repeated this procedure to attain at least 200 awakenings with a visual report for each participant. On average, we awakened participants every 342.0 s, and visual contents were reported in over 75% of the awakenings (Fig. 1B). Offline sleep stage scoring (fig. S2) further selected awakenings to exclude contamination from the wake stage in the period immediately before awakening (235, 198, and 186 awakenings for participants 1 to 3, respectively, used for decoding analyses) (*16*).

From the collected reports, words describing visual objects or scenes were manually extracted and mapped to WordNet, a lexical database in which semantically similar words are grouped as "synsets" in a hierarchical structure (*17*, *18*) (Fig. 2A). Using a semantic hierarchy, we grouped extracted visual words into base synsets that appeared in at least 10 reports from each participant (26, 18, and 16 synsets for participants 1 to 3, respectively; tables S2 to S4) (*16*). The fMRI data obtained before each awakening were labeled with a visual content vector, each element of which indicated the presence or absence of a base synset in the subsequent report (Fig. 2B and fig. S3). We also collected images depicting each base synset from ImageNet (*19*), an image database in which Web images are grouped according to WordNet, or from Google Images, for decoder training.

We constructed decoders by training linear support vector machines (SVMs) (*20*) on fMRI data measured while each person viewed Web images for each base synset. Multivoxel patterns in the higher visual cortex [HVC; the ventral region covering the lateral occipital complex (LOC), fusiform face area (FFA), and parahippocampal place area (PPA); 1000 voxels], the lower visual

[1]ATR Computational Neuroscience Laboratories, Kyoto 619-0288, Japan. [2]Nara Institute of Science and Technology, Nara 630-0192, Japan. [3]National Institute of Information and Communications Technology, Kyoto 619-0288, Japan.

*Present address: Brown University, 190 Thayer Street, Providence, RI 02912, USA.
†Present address: The University of Electro-Communications, Tokyo 182-8585, Japan.
‡Corresponding author. E-mail: kmtn@atr.jp

cortex (LVC; V1 to V3 combined; 1000 voxels), or the subareas (400 voxels for each area) were used as the input for the decoders (*16*).

First, a binary classifier was trained on the fMRI responses to stimulus images of two base synsets (three-volume averaged data corresponding to the 9-s stimulus block) and tested on the sleep samples [three-volume (9-s) averaged data immediately before awakening] that contained exclusively one of the two synsets while ignoring other concurrent synsets (*16*) (Fig. 3A). We only used synset pairs in which one of the synsets appeared in at least 10 reports without co-occurrence with the other (201, 118, and 86 pairs for participants 1 to 3, respectively). The distribution of the pairwise decoding accuracies for the HVC is shown together with that from the decoders trained on the same stimulus-induced fMRI data with randomly shuffled synset labels (Fig. 3B; fig. S4, individual participants). The mean decoding accuracy was 60.0%, 95% confidence interval (CI) [(59.0, 61.0%); three participants pooled], which was significantly higher than that of the label-shuffled decoders with both Wilcoxon rank-sum and permutation tests ($P < 0.001$).

To look into the commonality of brain activity between perception and sleep-onset imagery, we focused on the synset pairs that produced content-specific patterns in each of the stimulus and sleep experiments (pairs with high cross-validation classification accuracy within each of the stimulus and sleep data sets; figs. S5 and S6) (*16*). With the selected pairs, even higher accuracies were obtained [mean = 70.3%, CI (68.5, 72.1); Fig. 3B, dark blue; fig. S4, individual participants; tables S5 to S7, lists of the selected pairs], indicating that content-specific patterns are highly consistent between perception and sleep-onset imagery. The selection of synset pairs, which used knowledge of the test (sleep) data, does not bias the null distribution by the label-shuffled decoders (Fig. 3B, black), because the content specificity in the sleep data set alone does not imply commonality between the two data sets.

Additional analyses revealed that the multivoxel pattern, rather than the average activity level, was critical for decoding (figs. S7 and S8). We also found that the variability of decoding performance among synset pairs can be accounted for at least partly by the semantic differences between paired synsets. The decoding accuracy for synsets paired across meta-categories (human, object, scene, and others; tables S2 to S4) was significantly higher than that for synsets within meta-categories (Wilcoxon rank-sum test, $P < 0.001$; Fig. 3C and fig. S9). However, even within a meta-category, the mean decoding accuracy significantly exceeded chance level, indicating specificity to fine object categories.

The mean decoding accuracies for different visual areas are shown in Fig. 3D (fig. S10, individual participants). The LVC scored 54.3%, CI (53.4, 55.2) for all pairs, and 57.2%, CI (54.2, 60.2) for selected pairs (three participants pooled). The performance was significantly above chance level but worse than that for the HVC. Individual

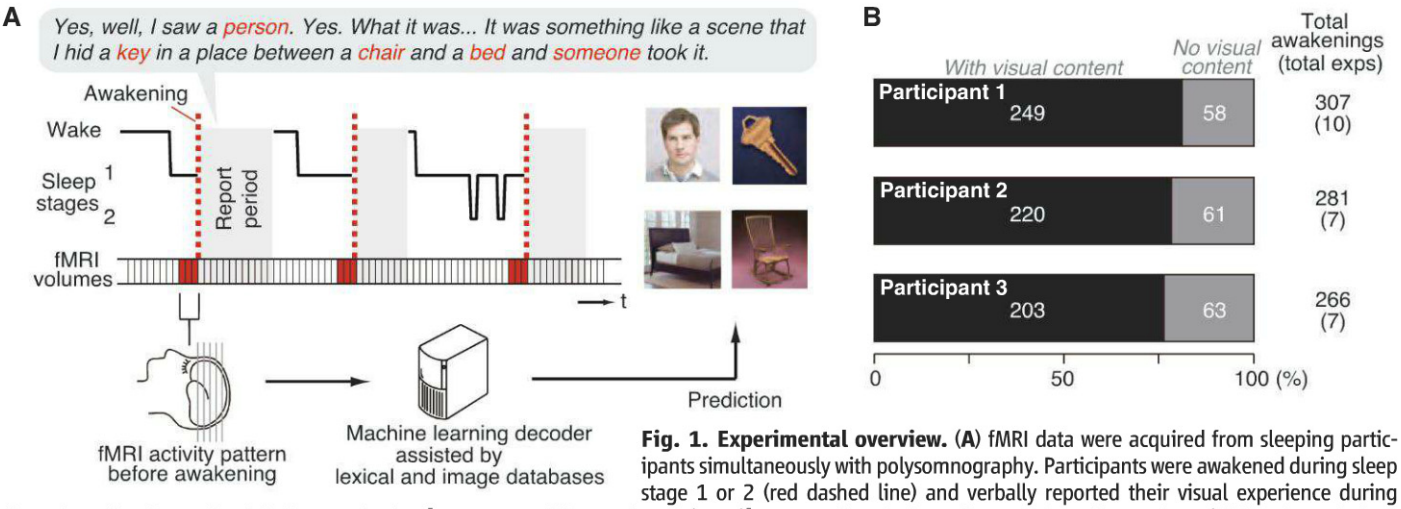


Fig. 1. Experimental overview. (**A**) fMRI data were acquired from sleeping participants simultaneously with polysomnography. Participants were awakened during sleep stage 1 or 2 (red dashed line) and verbally reported their visual experience during sleep. fMRI data immediately before awakening [an average of three volumes (= 9 s)] were used as the input for main decoding analyses (sliding time windows were used for time course analyses). Words describing visual objects or scenes (red letters) were extracted. The visual contents were predicted using machine-learning decoders trained on fMRI responses to natural images. (**B**) The numbers of awakenings with and without visual contents are shown for each participant (with numbers of experiments in parentheses).

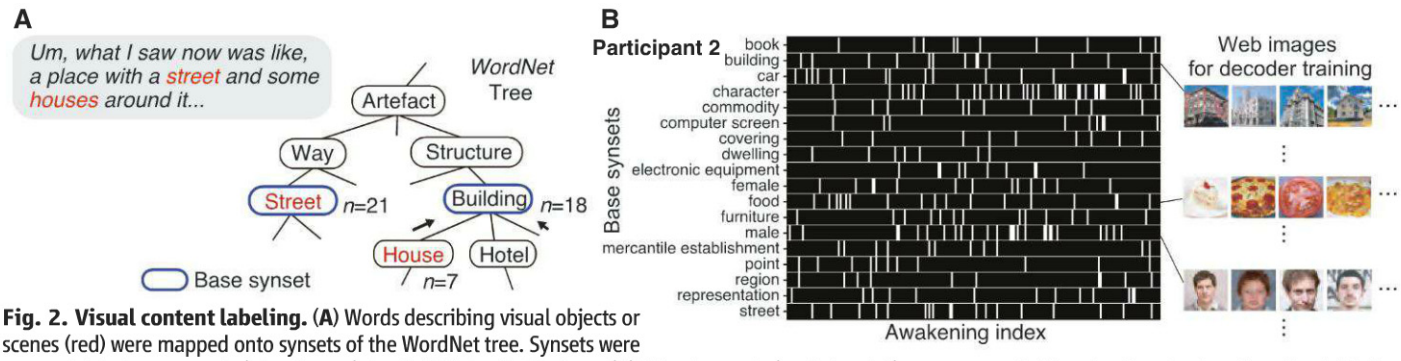


Fig. 2. Visual content labeling. (**A**) Words describing visual objects or scenes (red) were mapped onto synsets of the WordNet tree. Synsets were grouped into base synsets (blue frames) located higher in the tree. (**B**) Visual reports (participant 2) are represented by visual content vectors, in which the presence or absence of the base synsets in the report at each awakening is indicated by white or black, respectively. Examples of images used for decoder training are shown for some of the base synsets.

areas (V1 to V3, LOC, FFA, and PPA) showed a gradual increase in accuracy along the visual processing pathway, mirroring the progressively complex response properties from low-level image features to object-level features (*21*). When the time window was shifted, the decoding accuracy peaked around 0 to 10 s before awakening (Fig. 3E and fig. S11; no correction for hemodynamic delay). The high accuracies after awakening may be due to hemodynamic delay and the large time window. Thus, verbal reports are likely to reflect brain activity immediately before awakening.

To read out richer contents given arbitrary sleep data, we next performed a multilabel decoding analysis in which the presence or absence

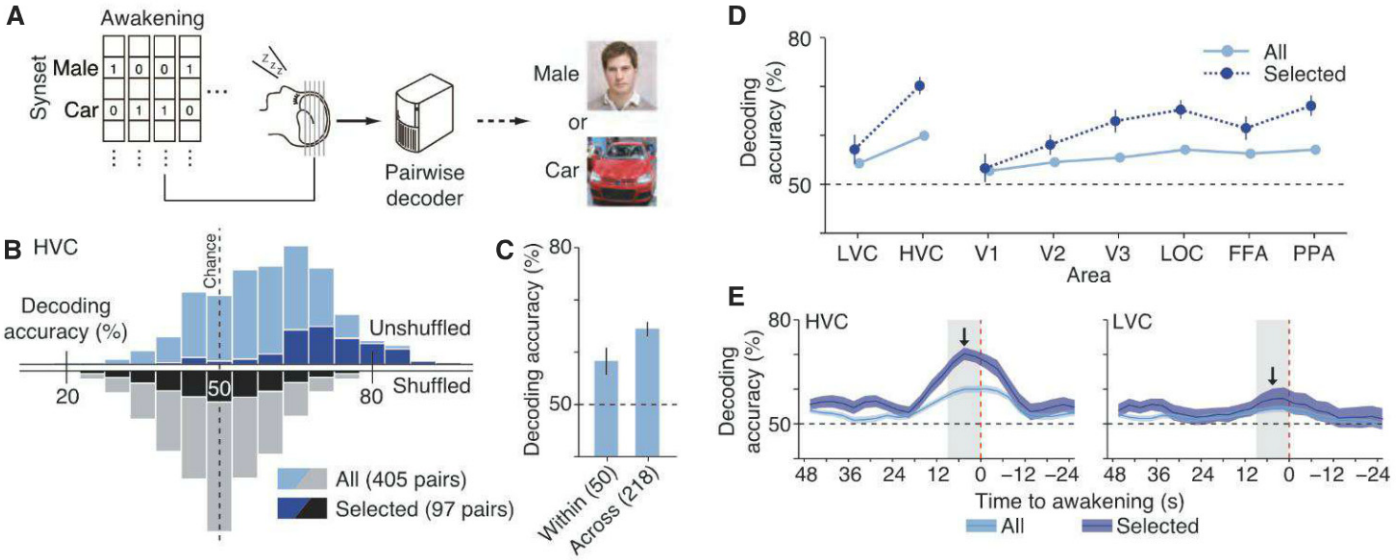


Fig. 3. Pairwise decoding. (**A**) Schematic overview. (**B**) Distributions of decoding accuracies with original and label-shuffled data for all pairs (light blue and gray) and selected pairs (dark blue and black) (three participants pooled). (**C**) Mean accuracies for the pairs within and across meta-categories (synsets in others were excluded; numbers of pairs are in parentheses). (**D**) Accuracies across visual areas (numbers of selected pairs for V1, V2, V3, LOC, FFA, PPA, LVC, and HVC: 45, 50, 55, 70, 48, 78, 55, and 97). (**E**) Time course (HVC and LVC; averaged across pairs and participants). The plot shows the performance with the 9-s (three-volume) time window centered at each point (gray window and arrow for main analyses). For all results, error bars or shadings indicate 95% CI, and dashed lines denote chance level.

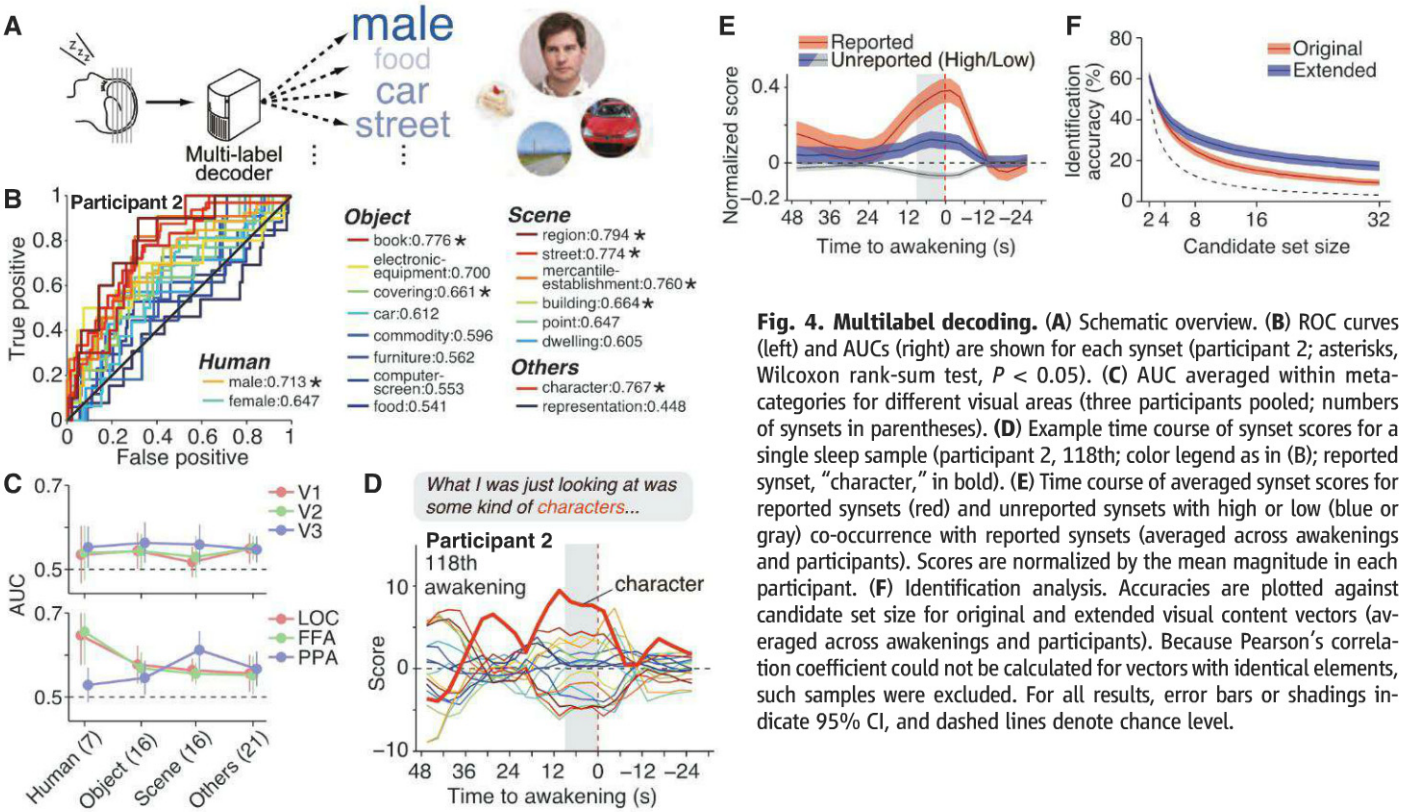


Fig. 4. Multilabel decoding. (**A**) Schematic overview. (**B**) ROC curves (left) and AUCs (right) are shown for each synset (participant 2; asterisks, Wilcoxon rank-sum test, $P < 0.05$). (**C**) AUC averaged within meta-categories for different visual areas (three participants pooled; numbers of synsets in parentheses). (**D**) Example time course of synset scores for a single sleep sample (participant 2, 118th; color legend as in (B); reported synset, "character," in bold). (**E**) Time course of averaged synset scores for reported synsets (red) and unreported synsets with high or low (blue or gray) co-occurrence with reported synsets (averaged across awakenings and participants). Scores are normalized by the mean magnitude in each participant. (**F**) Identification analysis. Accuracies are plotted against candidate set size for original and extended visual content vectors (averaged across awakenings and participants). Because Pearson's correlation coefficient could not be calculated for vectors with identical elements, such samples were excluded. For all results, error bars or shadings indicate 95% CI, and dashed lines denote chance level.

of each base synset was predicted by a synset detector constructed from a combination of pairwise decoders (Fig. 4A) (*16*). The synset detector provided a continuous score indicating how likely the synset is to be present in each report. We calculated receiver operating characteristic (ROC) curves for each base synset by shifting the detection threshold for the output score (Fig. 4B, the HVC in participant 2, time window immediately before awakening; fig. S12, all participants), and the detection performance was quantified by the area under the curve (AUC). Although the performance varied across synsets, 18 out of the total 60 synsets were detected with above-chance levels (Wilcoxon rank-sum test, uncorrected $P < 0.05$), greatly exceeding the number of synsets expected by chance ($0.05 \times 60 = 3$).

Using the AUC, we compared the decoding performance for individual synsets grouped into meta-categories in different visual areas. Overall, the performance was better in the HVC than in the LVC, consistent with the pairwise decoding performance [fig. S13; three participants pooled; analysis of variance (ANOVA), $P = 0.003$]. Although V1 to V3 did not show different performances across meta-categories, the higher visual areas showed a marked dependence on meta-categories (Fig. 4C and fig. S13). In particular, the FFA showed better performance with human synsets, whereas the PPA showed better performance with scene synsets [ANOVA (interaction), $P = 0.001$], consistent with the known response characteristics of these areas (*22*, *23*). The LOC and FFA showed similar results, presumably because our functional localizers selected partially overlapping voxels.

The output scores for individual synsets showed diverse and dynamic profiles in each sleep sample (Fig. 4D, fig. S14, and movies S1 and S2) (*16*). These profiles may reflect a dynamic variation of visual contents, including those experienced even before the period near awakening. On average, there was a general tendency for the scores for reported synsets to increase toward the time of awakening (Fig. 4E and fig. S15). Synsets that did not appear in reports showed greater scores if they had a high co-occurrence relationship with reported synsets (Fig. 4E; synsets with the top 15% conditional probabilities given a reported synset, calculated from the whole-content vectors in each participant). The effect of co-occurrence is rather independent of that of semantic similarity (Fig. 3C), because both factors (high/low co-occurrence and within/across meta-categories) had highly significant effects on the scores of unreported synsets (time window immediately before awakening; two-way ANOVA, $P < 0.001$, three participants pooled) with moderate interaction ($P = 0.016$). The scores for reported synsets were significantly higher than those for unreported synsets even within the same meta-category (Wilcoxon rank-sum test, $P < 0.001$). Verbal reports are unlikely to describe full details of visual experience during sleep, and it is possible that contents with high general co-occurrence (such as street and car) tend to be experienced together even when all are not reported. Therefore, high scores for the unreported synsets may indicate unreported but actual visual contents during sleep.

Finally, to explore the potential of multilabel decoding to distinguish numerous contents, we performed identification analysis (*7*, *8*). The output scores (score vector) were used to identify the true visual content vector among a variable number of candidates (true vector + random vectors with matched probabilities for each synset) by selecting the candidate most correlated with the score vector (repeated 100 times for each sleep sample to obtain the correct identification rate) (*16*). The performance exceeded chance level across all set sizes (Fig. 4F, HVC, three participants pooled; fig. S16, individual participants), although the accuracies were not as high as those achieved using stimulus-induced brain activity in previous studies (*7*, *8*). The same analysis was performed with extended visual content vectors in which unreported synsets having a high co-occurrence with reported synsets (top 15% conditional probability) were assumed to be present. The results showed that extended visual content vectors were better identified (Fig. 4F and fig. S16), suggesting that multilabel decoding outputs may represent both reported and unreported contents.

Together, our findings provide evidence that specific contents of visual experience during sleep are represented by, and can be read out from, visual cortical activity patterns shared with stimulus representation. Our approach extends previous research on the (re)activation of the brain during sleep (*24*–*27*) and the relationship between dreaming and brain activity (*2*, *3*, *28*) by discovering links between complex brain activity patterns and unstructured verbal reports using database-assisted machine-learning decoders. The results suggest that the principle of perceptual equivalence (*29*), which postulates a common neural substrate for perception and imagery, generalizes to spontaneously generated visual experience during sleep. Although we have demonstrated semantic decoding with the HVC, this does not rule out the possibility of decoding low-level features with the LVC. The decoding presented here is retrospective in nature: Decoders were constructed after sleep experiments based on the collected reports. However, because reported synsets largely overlap between the first and the last halves of the experiments (59 out of 60 base synsets appeared in both), the same decoders may apply to future sleep data. The similarity between REM and sleep-onset reports (*13*–*15*) and the visual cortical activation during the REM sleep (*24*, *25*, *28*) suggest that the same decoders could also be used to decode REM imagery. Our method may further work beyond the bounds of sleep stages and reportable experience to uncover the dynamics of spontaneous brain activity in association with stimulus representation. We expect that it will lead to a better understanding of the functions of dreaming and spontaneous neural events (*10*, *30*).

References and Notes

1. W. Dement, N. Kleitman, *J. Exp. Psychol.* **53**, 339 (1957).
2. M. Dresler *et al.*, *Curr. Biol.* **21**, 1833 (2011).
3. C. Marzano *et al.*, *J. Neurosci.* **31**, 6674 (2011).
4. J. V. Haxby *et al.*, *Science* **293**, 2425 (2001).
5. Y. Kamitani, F. Tong, *Nat. Neurosci.* **8**, 679 (2005).
6. S. A. Harrison, F. Tong, *Nature* **458**, 632 (2009).
7. K. N. Kay, T. Naselaris, R. J. Prenger, J. L. Gallant, *Nature* **452**, 352 (2008).
8. Y. Miyawaki *et al.*, *Neuron* **60**, 915 (2008).
9. M. Stokes, R. Thompson, R. Cusack, J. Duncan, *J. Neurosci.* **29**, 1565 (2009).
10. Y. Nir, G. Tononi, *Trends Cogn. Sci.* **14**, 88 (2010).
11. T. Hori, M. Hayashi, T. Morikawa, in *Sleep Onset: Normal and Abnormal Processes,* R. D. Ogilvie, J. R. Harsh, Eds. (American Psychological Association, Washington, DC, 1994), pp. 237–253.
12. R. Stickgold, A. Malia, D. Maguire, D. Roddenberry, M. O'Connor, *Science* **290**, 350 (2000).
13. D. Foulkes, G. Vogel, *J. Abnorm. Psychol.* **70**, 231 (1965).
14. G. W. Vogel, B. Barrowclough, D. D. Giesler, *Arch. Gen. Psychiatry* **26**, 449 (1972).
15. D. Oudiette *et al.*, *Conscious. Cogn.* **21**, 1129 (2012).
16. Materials and methods are available as supplementary materials on *Science* Online.
17. C. Fellbaum, Ed., *WordNet: An Electronic Lexical Database* (MIT Press, Cambridge, MA, 1998).
18. A. G. Huth, S. Nishimoto, A. T. Vu, J. L. Gallant, *Neuron* **76**, 1210 (2012).
19. J. Deng *et al.*, *IEEE Conference on Computer Vision and Pattern Recognition,* 248 (2009).
20. V. N. Vapnik, *Statistical Learning Theory* (Wiley, New York, 1998).
21. E. Kobatake, K. Tanaka, *J. Neurophysiol.* **71**, 856 (1994).
22. R. Epstein, N. Kanwisher, *Nature* **392**, 598 (1998).
23. N. Kanwisher, J. McDermott, M. M. Chun, *J. Neurosci.* **17**, 4302 (1997).
24. A. R. Braun *et al.*, *Science* **279**, 91 (1998).
25. P. Maquet, *J. Sleep Res.* **9**, 207 (2000).
26. Y. Yotsumoto *et al.*, *Curr. Biol.* **19**, 1278 (2009).
27. M. A. Wilson, B. L. McNaughton, *Science* **265**, 676 (1994).
28. S. Miyauchi, M. Misaki, S. Kan, T. Fukunaga, T. Koike, *Exp. Brain Res.* **192**, 657 (2009).
29. R. A. Finke, *Principles of Mental Imagery* (MIT Press, Cambridge, MA, 1989).
30. J. A. Hobson, *Nat. Rev. Neurosci.* **10**, 803 (2009).

Acknowledgments: We thank Y. Onuki, T. Beck, Y. Fujiwara, G. Pandey, and T. Kubo for assistance with early experiments and M. Takemiya and P. Sukhanov for comments on the manuscript. This work was supported by grants from the Strategic Research Program for Brain Science (MEXT), the Strategic Information and Communications R&D Promotion Programme (SOUMU), the National Institute of Information and Communications Technology, the Nissan Science Foundation, and the Ministry of Internal Affairs and Communications (Novel and innovative R&D making use of brain structures).

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1234330/DC1
Materials and Methods
Figs. S1 to S16
Tables S1 to S7
References (*31*–*43*)
Movies S1 and S2

20 December 2012; accepted 5 March 2013
Published online 4 April 2013;
10.1126/science.1234330

PCR INSTRUMENT

The new Mastercycler nexus X1 is the latest addition to the Mastercycler range of polymerase chain reaction (PCR) instruments. It provides researchers with increased heating and cooling rates by uniting the innovative software found in the Mastercycler nexus with a fast 96-well block. All this is achieved while maintaining low power consumption and low noise emission levels (≤ 40 dB[a]), making the Mastercycler nexus X1 suitable for work within busy life science laboratories. A heating rate of 5°C/s means run times using the Mastercycler nexus X1 are very short, allowing several users to work on it during the course of a day. Up to three units can be combined for maximum throughput, and a booking schedule integrated into the software allows easy allocation of time to fit researchers' schedules. When in use, the Mastercycler may be connected to a computer network, supplying e-mail status updates on the progress of the PCR run.

Eppendorf

For info: 800-645-3050 | www.eppendorfna.com

BISULFITE CONVERSION SYSTEM

The MethylEdge Bisulfite Conversion System offers an efficient method for rapidly performing bisulfite conversion and DNA clean-up within two hours, while maintaining DNA integrity to provide more intact, converted DNA for a variety of sensitive applications. MethylEdge requires only small amounts of input genomic DNA (100 pg to 2 μg). Using optimized conversion reagents, clean-up, and recovery methods, the system converts DNA at 90% efficiency with minimal fragmentation. The resulting DNA is suitable for use in downstream applications such as polymerase chain reaction, array, or sequencing assays. The kit includes all reagents necessary to perform 50 bisulfite conversion reactions. All MethylEdge reagents can be stored at room temperature and require minimal upfront preparation. The Promega MethylEdge Bisulfite Conversion System complements Promega nucleic acid purification kits for a wide variety of sample types and throughput levels, quantitation chemistries for both pre- and post-conversion, and Promega amplification technologies for detection.

Promega

For info: 608-274-4330 | www.promega.com/methylation

PCR THERMAL CYCLER

The FlexCycler^2 is a modern thermal cycler which combines exceptional design with reliable technology in one system. Using the Quick-X-Change block exchange system, block modules can be exchanged within seconds, allowing the instrument to flexibly adapt to changing requirements. In total, six different mono-and twin-block modules are available, all of which can replace each other. The two independent blocks of the twin-block modules allow two different polymerase chain reaction (PCR) programs to be run simultaneously, thereby helping users avoid bottlenecks. The 96-well block module and the 48-well twin-block module, for optimization of new primer pairs, are also an available option with gradient function. The FlexCycler^2 provides state-of-the-art heating and cooling rates, and excellent temperature uniformity enables reproducible conditions in all positions of the sample blocks. The user-friendly software, in combination with extensive additional software options, makes the FlexCycler^2 the perfect system for challenging PCR applications.

Analytik-Jena

For info: +49-36-41777-0 | www.analytik-jena.com

DNA LIBRARY QUANTIFICATION KIT

The new ddPCR Library Quantification Kit for Illumina TruSeq sample preparation kits enables QX100 Droplet Digital PCR system users to precisely and directly measure amplifiable library concentrations for next generation sequencing. The TruSeq sample preparation method is the technology behind Illumina's MiSeq and HiSeq next generation sequencing platforms. Using the ddPCR Library Quantification Kit to quantify TruSeq DNA libraries maximizes the number of useable reads, enables consistent loading, and optimizes the utilization of every sequencing run. The resulting data provides additional measures of library quality not provided by other methods, including the percentage of nonamplifiable species such as adapter dimers as well as the size range of library inserts.

Bio-Rad Laboratories

For info: 800-424-6723 | www.bio-rad.com

DNA REFERENCE STANDARD

The new Quantitative Multiplex DNA Reference Standard is the first of its kind and is intended for researchers assessing multiple biomarkers in a single assay using platforms such as next generation sequencing. As multiplexing assays and large tumor profiling projects become more common, standardization will be essential to enable confidence in experimental results. To date, a significant challenge has been access to reliable, renewable external reference standards. The novel Quantitative Multiplex DNA Reference Standard directly addresses this need by enabling researchers to quantify a range of detection thresholds for 11 cancer relevant mutations. This is accomplished across complex samples in a single assay in the form of renewable material originating from precisely engineered cell lines.

Horizon Diagnostics

For info: 800-860-3546 | www.horizondx.com

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/products/newproducts.dtl for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by Science or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

There's only one Science

Science Careers Advertising

For full advertising details, go to ScienceCareers.org and click For Employers, or call one of our representatives.

Tracy Holmes
Worldwide Associate Director
Science Careers
Phone: +44 (0) 1223 326525

THE AMERICAS

E-mail: advertise@sciencecareers.org
Fax: 202-289-6742

Tina Burks
East Coast/West Coast/South America
Phone: 202-326-6577

Marci Gallun
Midwest/Canada
Phone: 202-326-6582

Candice Nulsen
Corporate
Phone: 202-256-1528

Online Job Posting Questions
Phone: 202-312-6375

EUROPE / INDIA / AUSTRALIA / NEW ZEALAND / REST OF WORLD

E-mail: ads@science-int.co.uk
Fax: +44 (0) 1223 326532

Axel Gesatzki
Phone: +44 (0)1223 326529

Lucy Nelson
Phone: +44 (0)1223 326527

Kelly Grace
Phone: +44 (0) 1223 326528

JAPAN

Yuri Kobayashi
Phone: +81-(0)90-9110-1719
E-mail: ykobayas@aaas.org

CHINA / KOREA / SINGAPORE / TAIWAN / THAILAND

Ruolei Wu
Phone: +86-1367-1015-294
E-mail: rwu@aaas.org

All ads submitted for publication must comply with applicable U.S. and non-U.S. laws. *Science* reserves the right to refuse any advertisement at its sole discretion for any reason, including without limitation for offensive language or inappropriate content, and all advertising is subject to publisher approval. *Science* encourages our readers to alert us to any ads that they feel may be discriminatory or offensive.

Science Careers
From the journal *Science* AAAS

ScienceCareers.org

Plant Biotic Stress Biologist

IANR

The Institute of Agriculture and Natural Resources (IANR) at the University of Nebraska-Lincoln (UNL) is committed to world-class excellence in applications of agricultural and life sciences towards a sustained high quality of life for the citizens of Nebraska and a rapidly growing global population. Reflecting this commitment, IANR has announced an immediate initiative to hire more than 30 new tenure-track faculty members to fill strategic needs in impact areas of Science Literacy, Stress Biology of Plants and Animals, Healthy Humans, Healthy Agro/Eco Systems, and Computational Sciences. To view a listing of all of these positions, and for details as positions are released, please see **http://ianrhome.unl.edu/web/ianr/growingianr**.

Within this initiative, IANR is seeking applicants for an academic year tenure-leading plant biotic stress biologist position at the Assistant or Associate Professor level with a discovery research program focused on increasing our understanding of plant immunity at the molecular and cellular level and how plants prevent microbial diseases. The preference is that the research of the successful applicant be focused on the plant side of the host-microbe interaction with a model organism and/or a crop species. The tenure-home of this position will be in the Department of Plant Pathology with a joint appointment in the Center for Plant Science Innovation. The incumbent is expected to collaborate with faculty in the Stress Biology Initiative. A Ph.D. in plant biology, microbiology, plant pathology, or a closely related field is required and postdoctoral research experience is preferred. For consideration at the associate level the applicant must have a significant record of publications and external grant funding. The incumbent will be expected to teach courses in plant sciences or plant pathology, advise undergraduate and graduate students and contribute to curriculum development in the Plant Sciences.

To view the complete position details and make application for this position, go to the UNL Employment web site: **http://employment.unl.edu**. Search for requisition number **F_130119**. Click on "**Apply to this Job**". Complete the application and attach a letter of interest, an overview of research and teaching experience and interests, curriculum vitae, and contact information for three professional references. Review of applications will begin on **June 17, 2013** and continue until the position is filled or the search is closed.

The University of Nebraska has an active National Science Foundation ADVANCE gender equity program, and is committed to a pluralistic campus community through Affirmative Action, Equal Opportunity, work-life balance, and dual careers.

UNIVERSITY OF LIVERPOOL

Faculty of Health and Life Sciences
Institute of Translational Medicine
Department of Cellular and Molecular Physiology

2 Chairs in Cellular and Molecular Physiology

Salary Negotiable

The University of Liverpool invites applications for 2 academic posts within the Department of Cellular and Molecular Physiology.

For one post we are seeking a suitably skilled individual to drive development of state-of-the-art imaging platforms, including high field strength MRI, photoacoustics and next generation SPECT, as part of an initiative to establish a Centre for Preclinical Imaging. You should have the expertise and vision to create a world-class imaging resource serving both the local and national research communities. For the second post we are keen to appoint a senior academic to join our expanding programme of research in stem cell safety and/or the cancer microenvironment. You should have an outstanding track record of research, which complements on-going work in the Department, the MRC Centre for Drug Safety Science, or the Liverpool CRUK Centre.

Appointment at Reader/Senior Lecturer may also be considered.

Job Ref: A-579289/S **Closing Date: 31 May 2013**

For full details, or to request an application pack, visit www.liv.ac.uk/working/job_vacancies/ or e-mail jobs@liv.ac.uk
Please quote job ref in all enquiries.

COMMITTED TO DIVERSITY AND EQUALITY OF OPPORTUNITY

Stonewall DIVERSITY CHAMPION

POSITIVE ABOUT DISABLED PEOPLE

ROBERT J. AND CLAIRE PASAROW FOUNDATION
25th ANNUAL AND FINAL MEDICAL RESEARCH AWARDS
ANNOUNCEMENT OF AWARDEES
Cancer, Cardiovascular Disease, Neuropsychiatry

The Board of Directors of the Robert J. and Claire Pasarow Foundation is pleased to announce the recipients of the twenty-fifth Annual Medical Research Awards. The Foundation, created in 1987 by Mr. and Mrs. Pasarow of Beverly Hills, California, was established to celebrate extraordinary achievement, creativity, and distinction in the areas of cancer, cardiovascular disease, and neuropsychiatry. This year is especially significant as it will be the final year of what has proven to be a highly successful program.

The 25th Annual Medical Research Awardees:

Cancer Research: Elaine Fuchs, PhD (The Rockefeller University); **Professor Sir Richard Peto** (University of Oxford); and **Matthew P. Scott, PhD** (Stanford University)

Cardiovascular Disease Research: Hal Dietz, MD (Johns Hopkins University); **Charles T. Esmon, PhD** (University of Oklahoma); and **Helen H. Hobbs, MD** (University of Texas Southwestern)

Neuropsychiatry Research: Karl Deisseroth, MD, PhD (Stanford University); **Helen S. Mayberg, MD** (Emory University); and **Carla Shatz, PhD** (Stanford University)

Past Awardees of the Medical Research Awards:

Cancer Research (including basic cellular processes and the various forms of cancer): Peter K. Vogt, PhD; Irving L. Weissman, MD; George F. Vande Woude, PhD; Erkki Ruoslahti, MD; Harold N. Weintraub, MD, PhD; Ronald M. Evans, PhD; Stanley J. Korsmeyer, MD; Carlo M. Croce, MD; Alfred G. Knudson, Jr., MD, PhD; Robert A. Weinberg, MD; Eric S. Lander, D.Phil.; Paul L. Modrich, PhD; Anthony S. Fauci, MD; Alexander J. Varshavsky, PhD; Tom Maniatis, PhD; Roger D. Kornberg, PhD; Elizabeth H. Blackburn, PhD; Fred W. Alt, PhD; Bert O'Malley, MD; Tony Hunter, PhD; Bert Vogelstein, MD; Brian J. Druker, MD; Inder M. Verma, PhD; Angela H. Brodie, PhD; Lewis C. Cantley, PhD; Joan A. Steitz, PhD; David M. Livingston, MD; Joan Massagué, PhD; and Michael G. Rosenfeld, MD

Cardiovascular Disease Research (including disorders of the heart and vascular system): Burton E. Sobel, MD; Harvey Feigenbaum, MD; Bernardo Nadal-Ginard, MD, PhD; Mordecai P. Blaustein, MD; Jonathan Seidman, PhD and Christine Seidman, MD; Glenn A. Langer, MD; Philip Majerus, MD; Jan L. Breslow, MD; Kenneth R. Chien, MD, PhD; Michael A. Gimbrone, Jr., MD; Masashi Yanagisawa, MD, PhD; Mark T. Keating, MD; Eric N. Olson, PhD; Richard P. Lifton, MD, PhD; Robert J. Lefkowitz, MD; Shaun Coughlin, MD, PhD; Judah Folkman, MD; Barry S. Coller MD; Douglas C. Wallace, PhD; Daniel Steinberg, MD, PhD; Richard O. Hynes, PhD; David Ginsburg, MD; Jeffrey M. Friedman, MD, PhD; Andrew R. Marks, MD; Gerald M. Reaven, MD; Antonio M. Gotto, Jr., MD, D.Phil.; Robert W. Mahley, MD, PhD; and Peter J. Ratcliffe, MD

Neuropsychiatry Research (including neurologic and mental disorders): Nancy Wexler, PhD; Eric R. Kandel, MD; Floyd E. Bloom, MD; Solomon H. Snyder, MD; Michael E. Phelps, PhD; Patricia S. Goldman-Rakic, PhD; Huda Akil, PhD; Stanley Watson, MD, PhD; Arvid Carlsson, MD, PhD and Philip Seeman, MD, PhD; Stanley B. Prusiner, MD; Joseph T. Coyle, MD; Eric J. Nestler, MD, PhD; Fred H. Gage, PhD; Michael I. Posner, PhD and Marcus E. Raichle, MD; Pasko Rakic, MD, PhD; Seymour Benzer, PhD; Tomas Hökfelt, MD, PhD; Thomas M. Jessell, PhD; Judith L. Rapoport, MD; Bruce McEwen, PhD; Huda Y. Zoghbi, MD; Aaron T. Beck, MD; Jean-Pierre Changeux, PhD; Robert C. Malenka, MD, PhD; Roger A. Nicoll, MD; Charles F. Stevens, MD, PhD; Christine Petit, MD, PhD; Virginia M-Y Lee, PhD; and John Q. Trojanowski, MD, PhD

The Foundation's Board would like to express their appreciation to the many medical leaders, including deans of medical schools, chairs of departments, heads of research institutions, and past awardees, for their invaluable participation in the nomination and evaluation processes over the years. Board members: ***Michael R. Pasarow****, Chairman;* ***Anthony H. Pasarow****, Co-Treasurer;* ***Susan A. Pasarow, MSW****, Co-Treasurer;* ***Jack D. Barchas, MD*** *– Weill Cornell; Foundation President;* ***Shaun Coughlin, MD, PhD*** *– University of California, San Francisco;* ***Christopher J. Evans, PhD*** *– Brain Research Institute of UCLA;* ***Ronald M. Evans, PhD*** *– The Salk Institute;* ***Brian E. Henderson, MD*** *– University of Southern California;* ***Alexander J. Varshavsky, PhD*** *– CalTech, and* ***Joseph P. Van Der Meulen, MD*** *– University of Southern California (Emeritus).*

We are, of course, forever grateful for the generosity and vision of ***Robert and Claire Pasarow*** *who, throughout their lives, served the Foundation with enthusiasm and joy and relished in the entrepreneurial spirit of the awardees.*

The EGL Charitable Foundation invites you to apply to the

Gruss Lipper Post-Doctoral Fellowship Program

Eligibility

- Israeli citizenship
- Candidates must have completed PhD and/or MD/PhD degrees in the Biomedical Sciences at an accredited Israeli University/Medical School or be in their final year of study
- Candidates must have been awarded a postdoctoral position in the U.S. host research institution.

Details regarding the fellowship are available at www.eglcf.org

NW210508R

Job Announcements for Faculty Positions In Children's Nutrition Research

The Children's Nutrition Research Center of the Baylor College of Medicine announces the availability of TENURE-TRACK POSITIONS in nutrition at the behavior, molecular, cellular, and whole organism levels, including translational human nutrition research. Although positions are not limited to these areas, specific interests are fetal, neonatal, and early childhood nutrition (including fundamental mechanisms for the developmental origins of adult health and disease) and obesity (including fundamental adipose tissue biology, regulation of appetite and satiety, development of eating and physical activity behaviors, and prevention/intervention studies).

The level of appointment will be consistent with the candidate's qualifications. Candidates should have a record of achievement through publications in peer-review journals and successful competition for extramural funding. In addition to a strong independent research program, the successful candidate will be expected to have significant potential for research collaboration with existing research programs within the Children's Nutrition Research Center.

Interested candidates should submit their curriculum vitae, a statement of research interests, and three letters of recommendation to: **Dennis M. Bier, M.D., Director, Children's Nutrition Research Center, Baylor College of Medicine, 1100 Bates, Houston, TX 77030-2600**.

Baylor College of Medicine is an Equal Access/Equal Opportunity/ Affirmative Action Employer.

Assistant/Associate Professor of Immunology

The Ragon Institute of MGH, MIT and Harvard seeks outstanding scientists to join the faculty at the level of Assistant Professor/Associate Professor.

Candidates should demonstrate a record of outstanding achievement and potential to develop an independent research program focused on the ultimate goal of harnessing the immune system to prevent and cure human diseases. Candidates that bring complementary approaches for HIV vaccine development from engineering and physical sciences as well as those with virology and immunology expertise are encouraged to apply.

The successful candidate will have a M.D., Ph.D., M.D. /Ph.D. or equivalent degree, and build and maintain an internationally recognized extramurally funded research program. We particularly seek individuals whose expertise complements existing programs at the Ragon Institute, who possess the ability to work collaboratively with other scientists, and the scholarly qualities required to mentor doctoral students in the graduate programs. Academic rank will be determined in accordance with the successful candidate's experience and productivity.

Candidates should submit a cover letter, curriculum vitae including list of publications, and a brief statement of research interest by email (**Subject line: Job ID 2240687**) to: **ragoninstituterecruitment@partners.org**

Directly submit three letters of reference to:
Ragon Institute Search Committee (Job ID 2240687)
Ragon Institute of MGH, MIT and Harvard
400 Technology Square, Room 874
Cambridge, MA 02139

Please see **www.ragoninstitute.org/about/positions** for more information on this position.

Submission deadline: June 1, 2013

Massachusetts General Hospital and Harvard Medical School are Equal Opportunity Employers.

Faculty Position in Quantitative Biology

The **Abdus Salam International Centre for Theoretical Physics (ICTP)** seeks applications for a faculty position in quantitative biology starting from 1 October 2013.

ICTP, based in Trieste, Italy, provides a top class multi-disciplinary research environment in theoretical sciences, with cutting edge research, education and training, and it has been a driving force behind global efforts for the advancement of science in the developing world. We are now seeking to establish a new research section in quantitative biology in order to respond to the theoretical challenges arising from recent advances in large-scale experiments in biology and neurosciences. The programme will be open to, and actively seeking collaborations with, external partners.

Areas of interest include: biological networks; genomics, transcriptomics and proteomics; bioinformatics, machine learning and statistical inference approaches to biological systems; population genetics and evolution biology; and systems and computational neuroscience.

The successful applicant will play a leading role in establishing the new section, promoting cutting-edge research in quantitative biology, and graduate student supervision. Applicants should have an outstanding international reputation and a solid track record in obtaining external funding.

The position will be at the P3 UNESCO level, with a starting salary of $89200. Other benefits include family allowance, home travel, education grant for dependent children, pension plan and medical insurance. Applicants should submit their application through the UNESCO website (**https://unesco.taleo.net/careersection/2/joblist.ftl**) before **4 June 2013**.

For further information please contact qbio@ictp.it.

SUNNYBROOK PRIZE

RECOGNIZING EXCELLENCE IN UNDERGRADUATE RESEARCH

$10,000 PRIZE AWARDED TO THE WINNER

The competition recognizes excellence in undergraduate research and 3rd and 4th year undergraduate students are invited to submit their applications.

2014

Eligibility

- Should be in your 3rd or 4th year of study at a Canadian University by Fall 2013
- Completion of a research project with a focus on biomedical research

Application Requirements

- Cover letter
- Academic transcript
- Curriculum Vitae
- Two (2) reference letters
- Two (2) page abstract

2013 Winner
Ben Ouyang

2012 Winner
Chi-Han Chang

Formatting Requirements

- 2cm page margin
- 12 point font black ink
- Six (6) lines per inch – no condensed typing or spacing
- Print must be of letter quality and easy to read
- Diagrams and images will be counted within the 2 page limit

Finalists will be invited to present their research at the Sunnybrook Prize Event on Friday 3rd January, 2014

E-mail application packages to
marcelline.ramcharan@sri.utoronto.ca
Subject Line: Sunnybrook Prize 2014

SUBMISSION DEADLINE
THURSDAY 12TH SEPTEMBER, 2013

Sunnybrook RESEARCH INSTITUTE

Affiliated with the UNIVERSITY OF TORONTO

Science Careers online @sciencecareers.org

Science Careers

There's only one

Dr. Shirley Malcom

To Dr. Shirley Malcom, born and raised in the segregated South more than 65 years ago, a career based on her studies in science seemed even less likely than the launch of the Soviet's Sputnik. But with Sputnik's success, the Space Race officially started and, in an instant, brought a laser-like focus to science education and ways to deliver a proper response. Not long after, Dr. Malcom entered the picture.

Although black schools at the time received fewer dollars per student and did not have sufficient resources to maintain their labs at a level equivalent to the white schools, Dr. Malcom found her way to the University of Washington where she succeeded in obtaining a B.S. in spite of the difficulties of being an African American woman in the field of science. From there she went on to earn a Ph.D. in ecology from Penn State and held a faculty position at the University of North Carolina, Wilmington.

Dr. Malcom has served at the AAAS in multiple capacities, and is presently Head of the Directorate for Education and Human Resources Programs. Nominated by President Clinton to the National Science Board, she also held a position on his Committee of Advisors on Science and Technology. She is currently a member of the Caltech Board of Trustees, a Regent of Morgan State University, and co-chair of the Gender Advisory Board of the UN Commission on Science and Technology for Development. She has held numerous other positions of distinction and is the principal author of *The Double Bind: The Price of Being a Minority Woman in Science*.

Of her active career in science, Dr. Malcom says, "I guess I have become a poster child for taking one's science background and using that in many other ways: we ask questions; we try to understand what we find; we consider what evidence we would need to confirm or refute hypotheses. And that happens in whatever setting one finds oneself."

At *Science* we are here to help you in your own scientific career with expert career advice, forums, job postings, and more — all for free. Visit *Science* today at **ScienceCareers.org**.

For your career in science, there's only one Science

AAAS

ScienceCareers.org

Career advice | Job postings | Job Alerts | Career Forum | Crafting resumes/CVs | Preparing for interviews

Early Career Fellowships

For promising newly qualified postdoctoral researchers to launch their independent research career in India

The 4 year Fellowship support provides:
- Competitive personal support
- Generous Research funds
- Funding to work overseas for up to 2 years and develop international collaborations

Eligibility :

For Applicants in the last year of PhD or with no more than 4 years of Postdoctoral research experience.
For more information please visit
http://wellcomedbt.org/early.html

We encourage Clinicians and Public Health Researchers to apply in a separate competition to be launched by this year end.
For more information please visit
http://wellcomedbt.org/clinical_researcher.html

Preliminary Application due by 13 June, 2013.

Images courtesy Wellcome Images

wellcome trust DBT
INDIA ALLIANCE
The Wellcome Trust/DBT India Alliance is a registered charity in India

For further information and the preliminary application form
Please visit our website www.wellcomedbt.org

It's not because we all wear white coats that we all do the same job

INRA, ACCESSIBLE TO ALL AMBITIONS

As part of its Human Resources policy, INRA recruits research scientists and engineers with disabilities: these positions are within strategic areas of the Institute. You are invited to join our research units focused on the scientific challenges of tomorrow.
To apply, please download the application form available on our website **job.inra.fr/en**

orc.fr - Photo credits: Corbis

Applications until June 14th 2013
on **jobs.inra.fr/en**

INRA
SCIENCE & IMPACT

COMPLEX SYSTEMS SUMMER SCHOOL CHILE

Sponsored by the **Santa Fe Institute** and **Universidad del Desarrollo**

Chile,
November 11-21, 2013

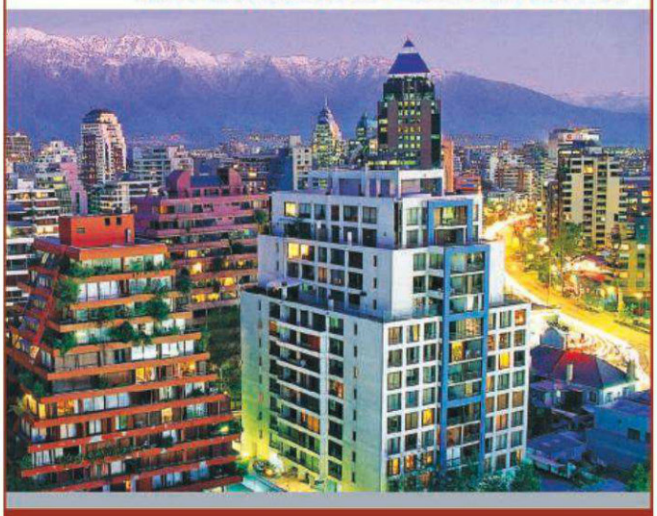
CSSS-Chile is an intensive 11-day study of biocomplexity and behavior in social and environmental complex systems, for graduate and postdoctoral researchers.

This program explores today's sustainability and social dynamics challenges through scaling, network theory, ecology, nonlinear dynamics, and other perspectives from complex systems science.

No tuition is charged. Accommodations and meals are provided. Support for travel may also be available.

Apply by July 15, 2013

To apply and for more information,
www.santafe.edu/chile



SANTA FE INSTITUTE

Universidad del Desarrollo
Universidad de Excelencia

POSITIONS OPEN

Beth Israel Deaconess Medical Center

JUNIOR- TO MID-LEVEL SCIENTIST

The Department of Radiation Oncology at the Beth Israel Deaconess Medical Center is recruiting a junior- to mid-level scientist to study immune response to ionizing radiation as part of a departmental program to exploit radioimmunotherapy in the clinic. Preference will be given to applicants with a strong background in tumor immunology and a record of significant external funding. Successful applicants will receive an appointment at Harvard Medical School. Applications, containing three references and curriculum vitae should be sent to: **Mary Parkman, assistant to Dr. Mary Ann Stevenson, Department of Radiation Oncology, 300 Brookine Avenue, Boston, MA 02215; e-mail: mparkman@bidmc.harvard.edu.**

MOLECULAR TUMOR BIOLOGY

A **POSTDOCTORAL** position is available in the laboratory of **Dr. Usha Kasid** at the Georgetown University Medical Center, Washington, DC to study novel molecular targets, signaling pathways and gene networks in cancer cells and to develop therapeutic strategies in animal models. Recent Ph.D. with strong background/training and hands-on experience in molecular biology and biochemistry is desirable. Understanding of the system biology and bioinformatics approaches for analyzing data from high-resolution genome-wide techniques is a plus. Please send curriculum vitae and the names, telephone numbers, and e-mail addresses of three references to **e-mail: kasidu@georgetown.edu.** *An Equal Opportunity/Affirmative Action Employer.*

FACULTY POSITION
University of Pittsburgh

The School of Pharmacy is seeking a creative faculty member who will become part of a team that believes in personalizing the education of the next generation of pharmacists and pharmaceutical scientists. The incumbent will be expected to contribute to courses in anatomy and physiology as well as biochemistry and to therapeutic modules (pharmacology and therapeutics, **website: http://www.pharmacy.pitt.edu/**). Scholarship is expected and could focus on the scholarship of teaching or involve research or clinical programs of the School of Pharmacy. Experience with or interest in online or blended synchronous and asynchronous learning is highly desirable as well as the interest in assessing the effectiveness of teaching and learning. Applicants should have a Ph.D., PharmD, M.D., or equivalent and be committed to providing innovative and dedicated teaching. The position, which is outside the tenure stream, can be a nine-month or calendar-year appointment depending on the applicant. Faculty rank and salary will be commensurate with experience and qualifications.

The School of Pharmacy is committed to providing students with a personalized education that will efficiently prepare them to innovate, lead, and identify opportunities to improve health using the clinical and research principles of the pharmaceutical sciences.

Applicants should send a letter describing their interest in the position, their curriculum vitae, and the names of at least four individuals who will serve as references to **Ms. Michele Chamberlain** at **e-mail: mrc6@pitt.edu.** *The University of Pittsburgh is an Affirmative Action/Equal Opportunity Employer.*

☑ More scientists agree —we are the most useful website.
www.ScienceCareers.org

POSITIONS OPEN

Yale University School of Medicine

RESEARCH ASSOCIATE in Ion Channel Electrophysiology

Position available for experienced patch-clamp electrophysiologist to join ongoing multidisciplinary research program exploring molecular genetic and biophysical basis for painful neuropathies. M.D. and/or Ph.D. degree, and experience in voltage-clamp electrophysiology with publications, are essential. Experience with analysis of DRG neurons, molecular analysis of Na channels and/or current-clamp electrophysiology are highly desirable. See: **Dib-Hajj** *et al., Nature Reviews Neuroscience* **14:** 49–62, 2013. Superb opportunity for a skilled electrophysiologist to work collaboratively with molecular geneticists, molecular and cell biologists, physiologists, and pharmacologists in a highly collaborative, interactive program. Send letter of interest, curriculum vitae, and three letters of reference to **Stephen G. Waxman,** M.D., Ph.D. **e-mail: stephen.waxman@yale.edu.**

MECHANICAL ENGINEERING

The Department of Mechanical Engineering at Iowa State University (ISU) invites applications for multiple tenure-track faculty positions at the **ASSISTANT, ASSOCIATE,** or **FULL PROFESSOR** ranks to begin in 2014 (**website: http://www.me.iastate.edu**). Exceptional candidates in all areas of mechanical engineering will be considered, with particular emphasis in: Manufacturing (additive, soft materials, micro/nano, high-speed, energy systems); Controls (robotics, complex systems, energy, biological, environmental); Biotechnology (biosensors and biochips, bioM/NEMS, biomaterials); Energy (building efficiency, systems engineering, renewable); Complex fluids (chemically reacting, biofluids, multiscale); Design (lifecycle analysis, sustainability, modeling and visualization). All interested, qualified persons are encouraged to apply early at **website: http://www.iastatejobs.com/applicants/Central?quickFind=83186** by completing the Employment Application for **vacancy #130200,** with applications to be reviewed on a continuing basis from July 1 to December 15, 2013. *ISU is an Equal Opportunity/Affirmative Action Employer, and we are seeking candidates who share this mission of advancing diversity.*

Help employers find you. Post your resume/cv.

Science Careers
From the journal *Science* AAAS
www.ScienceCareers.org

Find your future here.
www.ScienceCareers.org

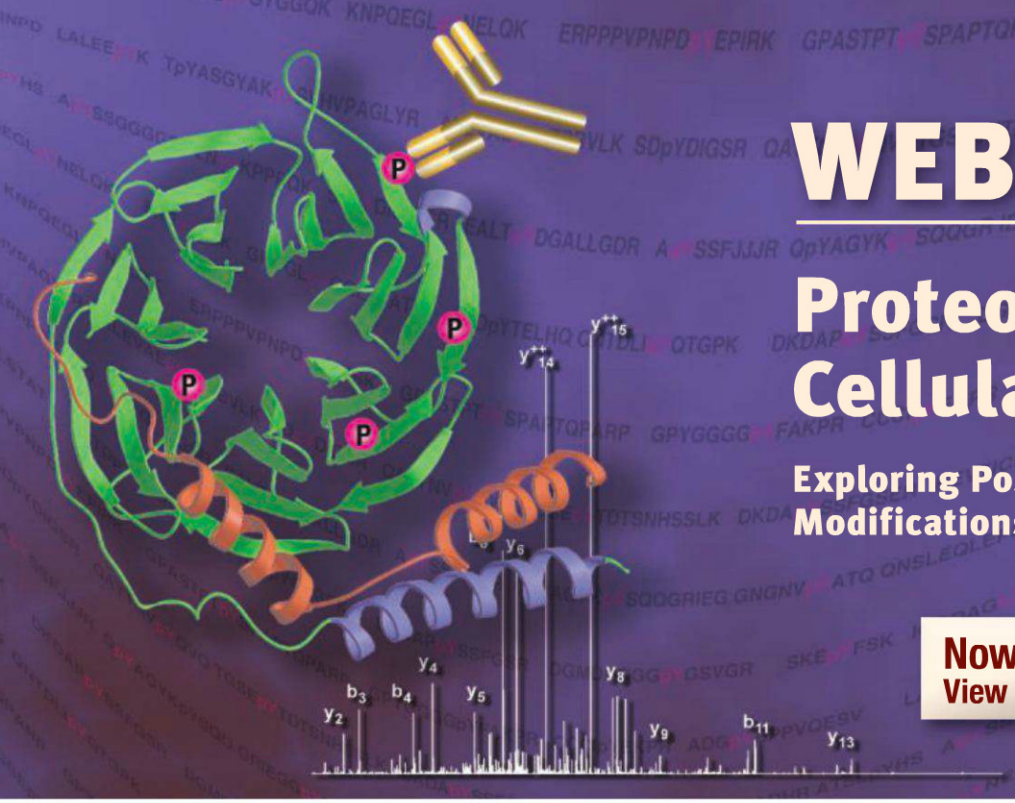
Nontraditional Careers: Opportunities Away From the Bench Webinar

Want to learn more about exciting and rewarding careers outside of academic/industrial research? View a roundtable discussion that looks at the various career options open to scientists and strategies you can use to pursue a nonresearch career.

Now Available On Demand
www.sciencecareers.org/webinar

Produced by the *Science*/AAAS Business Office.

Science Careers
From the journal *Science* AAAS



WEBINAR

Proteomics Meets Cellular Signaling

Exploring Post-Translational Modifications by Mass Spectrometry

Now Available On Demand!
View at webinar.sciencemag.org

Recorded on: Wednesday, May 1, 2013

12 noon Eastern, 9 a.m. Pacific, 5 p.m. UK, 6 p.m. CEST

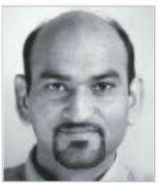
Post-translational modifications (PTMs)—including phosphorylation, acetylation, and ubiquitination—are specific modifications that can alter the activity of an individual protein target. The cumulative effect of these small modifications is the regulation of large signaling pathways and networks within cells.

Mass spectrometry is a powerful tool for the identification of individual PTMs during the investigation of signaling networks. A new approach using antibody-based peptide enrichment for a class of PTMs combined with tandem mass spectrometry has been successfully used for global identification of hundreds to thousands of PTMs within a sample. In this webinar, examples of how the technology was applied to each of these PTM classes will be described.

During this webinar, viewers will:

- Learn about post-translational modifications and their role in cellular signaling
- Discover how antibody-based proteomics approaches can be applied to identify, study, and characterize known and novel modifications
- Hear specific examples of how these approaches are being applied to the dissection of cellular signaling pathways, target identification, validation, and biomarker discovery by industry and academic experts

Speakers



Chunaram Choudhary, Ph.D.,
The Novo Nordisk Foundation Center for Protein Research

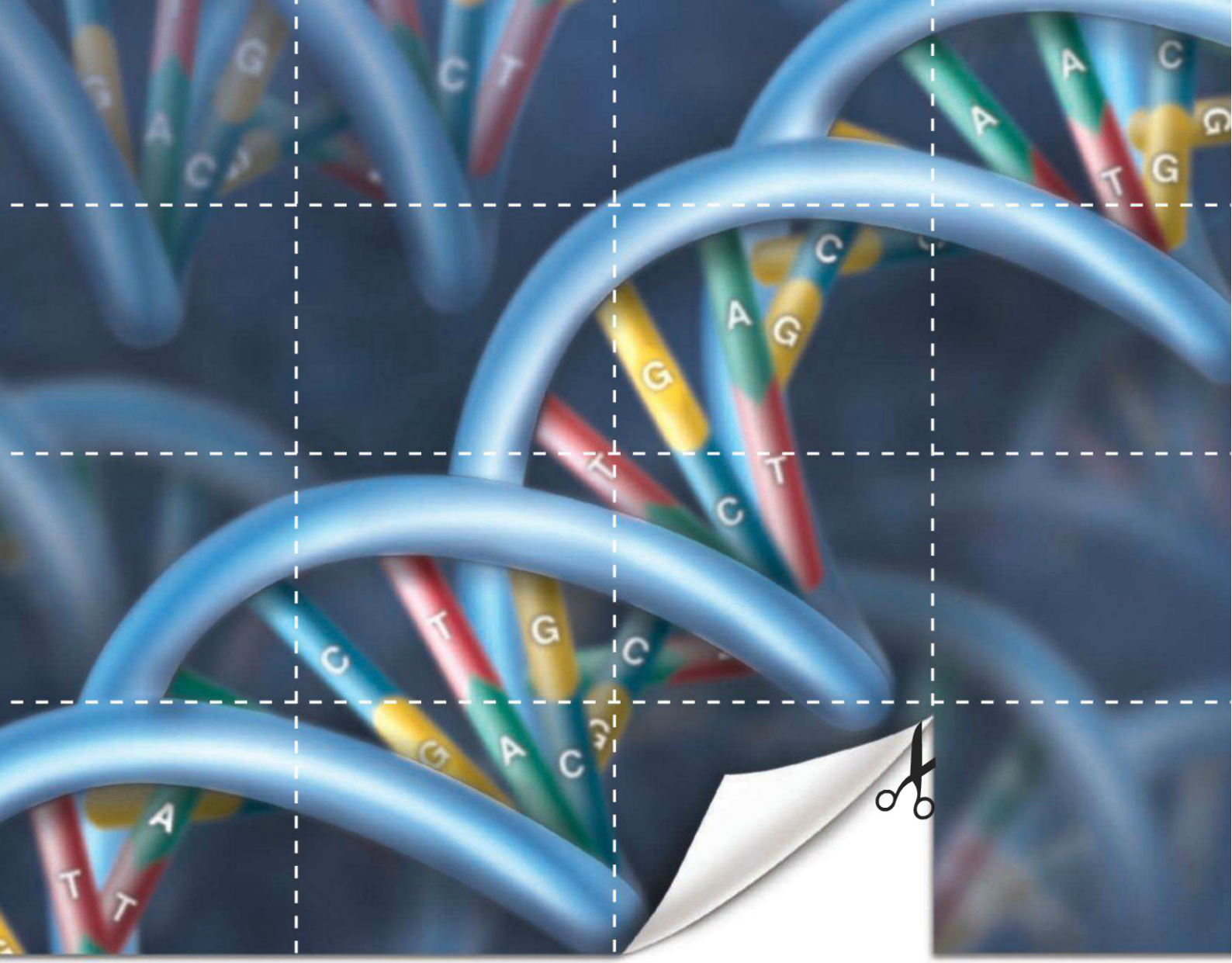
Cloud Paweletz, Ph.D.,
Dana Farber Cancer Institute

Vipin Suri, Ph.D.,
Sirtris, a GSK Company

Brought to you by the *Science*/AAAS Custom Publishing Office

Webinar sponsored by

Cell Signaling TECHNOLOGY®



Cut smarter.

Restriction enzymes from NEB – now with CutSmart™ Buffer

You make smart choices every day. Why stop there? The choice to use restriction enzymes from New England Biolabs is now even easier.

- Choose from > 200 restriction enzymes supplied with a single buffer
- Simplify your double digest reactions
- Reduce your pipetting steps by no longer having to add BSA

Now, that's **Smart!**

Explore the smarter choice at
NEBCutSmart.com

NEW ENGLAND BioLabs® Inc.
enabling technologies in the life sciences